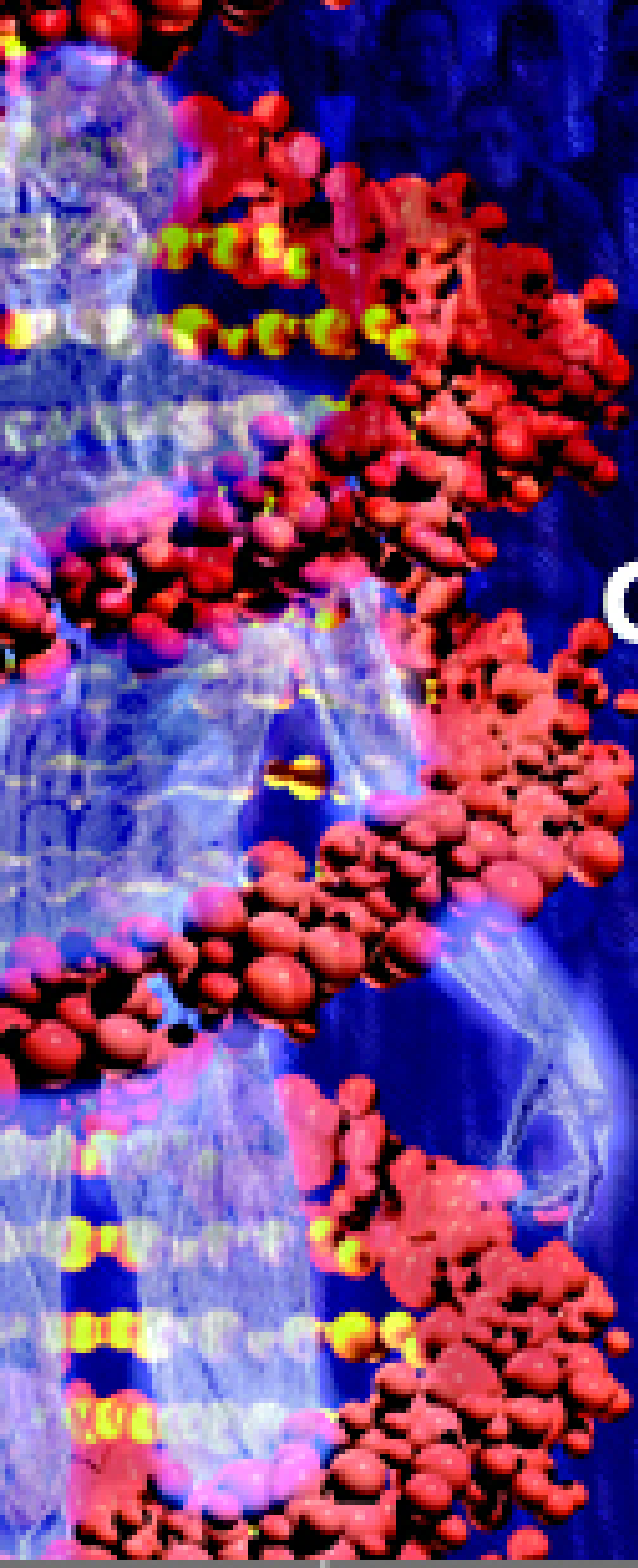

Edited by
Alan Wright
and Nicholas Hastie

Foreword by David J. Weatherall

Genes and Common Diseases

Genetics in Modern Medicine

CAMBRIDGE
Medicine

CAMBRIDGE www.cambridge.org/9780521833394

VISIT…

Genes and Common Diseases

Genes and common diseases presents an up-to-date view of the role of genetics in modern medicine, reflecting the strengths and limitations of a genetic perspective.

The current shift in emphasis from the study of rare single gene disorders to common diseases brings genetics into every aspect of modern medicine, from infectious diseases to therapeutics. However, it is unclear whether this increasingly genetic focus will prove useful in the face of major environmental influences in many common diseases.

The book takes a hard and self-critical look at what can and cannot be achieved using a genetic approach and what is known about genetic and environmental mechanisms in a variety of common diseases. It seeks to clarify the goals of human genetic research by providing state-of-the art insights into known molecular mechanisms underlying common disease processes while at the same time providing a realistic overview of the expected genetic and psychological complexity.

Alan Wright is a Programme Leader at the MRC Human Genetics Unit in Edinburgh.

Nicholas Hastie is Director of the MRC Human Genetics Unit in Edinburgh.

Genes and Common Diseases

Alan Wright

Nicholas Hastie

Foreword by David J. Weatherall

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK
Published in the United States of America by Cambridge University Press, New York
www.cambridge.org
Information on this title: www.cambridge.org/9780521833394

© Cambridge University Press 2007

This publication is in copyright. Subject to statutory exception and to the provision of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published in print format 2007

ISBN-13 978-0-511-33531-0 eBook (NetLibrary)
ISBN-10 0-511-33531-8 eBook (NetLibrary)

ISBN-13 978-0-521-83339-4 hardback
ISBN-10 0-521-83339-6 hardback

ISBN-13 978-0-521-54100-8 paperback
ISBN-10 0-521-54100-X paperback

Cambridge University Press has no responsibility for the persistence or accuracy of urls for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

Contents

List of Contributors *page* vii
Foreword xiii

Section 1: Introductory Principles

1 Genes and their expression 3
Dirk-Jan Kleinjan

2 Epigenetic modification of chromatin 20
Donncha Dunican, Sari Pennings and Richard Meehan

3 Population genetics and disease 44
Donald F. Conrad and Jonathan K. Pritchard

4 Mapping common disease genes 59
Naomi R. Wray and Peter M. Visscher

5 Population diversity, genomes and disease 80
Gianpiero L. Cavalleri and David B. Goldstein

6 Study design in mapping complex disease traits 92
Harry Campbell and Igor Rudan

7 Diseases of protein misfolding 113
Christopher M. Dobson

8 Aging and disease 132
Thomas T. Perls

9 The MHC paradigm: genetic variation and complex disease 142
Adrian P. Kelly and John Trowsdale

10 Lessons from single gene disorders 152
Nicholas D. Hastie

11 Environment and disease 164
A. J. McMichael and K. B. G. Dear

12 Contemporary ethico-legal issues in genetics 176
Renate Gertz, Shawn Harmon and Geoffrey Pradella

Section 2: Common Medical Disorders

13 Developmental disorders 201
Stephen P. Robertson and Andrew O. M. Wilkie

14 Genes, environment and cancer 213
D. Timothy Bishop

15 The polygenic basis of breast cancer 224
Paul D. P. Pharoah and Bruce A. J. Ponder

16 *TP53*: A master gene in normal and tumor suppression 233
Pierre Hainaut

17 Genetics of colorectal cancer 245
Susan M. Farrington and Malcolm G. Dunlop

18 Genetics of autoimmune disease 268
John I. Bell and Lars Fugger

19 Susceptibility to infectious diseases 277
Andrew J. Walley and Adrian V. S. Hill

20 Inflammatory bowel diseases 302
Jean-Pierre Hugot

21 Genetic anemias 316
W. G. Wood and D. R. Higgs

22 Genetics of chronic disease: obesity 328
I. Sadaf Farooqi and Stephen O'Rahilly

23 Type 2 diabetes mellitus 344
Mark I. McCarthy

24 Genetics of coronary heart disease 359
Rossi Naoumova, Stuart A. Cook, Paul Cook and Timothy J. Aitman

25 Genetics of hypertension 377
B. Keavney and M. Lathrop

26 Obstructive pulmonary disease 391
Bipen D. Patel and David A. Lomas

27 Skeletal disorders 406
Robert A. Colbert

28 The genetics of common skin diseases 427
Jonathan Rees

29 Molecular genetics of Alzheimer's disease and other adult-onset dementias 439
P. H. St George-Hyslop

30 Major psychiatric disorders in adult life 454
Amanda Elkin, Sridevi Kalidindi, Kopal Tandon and Peter McGuffin

31 Speech and language disorders 469
Gabrielle Barnby and Anthony J. Monaco

32 Common forms of visual handicap 488
Alan Wright

33 Genetic and environmental influences on hearing impairment 505
Karen P. Steel

34 Pharmacogenomics: clinical applications 516
Gillian Smith, Mark Chamberlain and C. Roland Wolf

Index 529

Contributors

Adrian V. S. Hill
Human Genetics
University of Oxford
Wellcome Trust Centre for
Human Genetics
Oxford, UK

Adrian P. Kelly
Immunology Division
Department of Pathology
Cambridge, UK

A. J. McMichael
National Centre for Epidemiology and
Population Health
The Australian National University
Canberra, Australia

Alan Wright
MRC Human Genetics Unit
Western General Hospital
Edinburgh, UK

Amanda Elkin
Neurogenetics Group
Wellcome Trust Centre for Human Genetics
Oxford, UK

Andrew J. Walley
Complex Human Genetics
Imperial College London
Section of Genomic Medicine
Hammersmith Hospital
London, UK

Andrew O.M. Wilkie
Weatherall Institute of
Molecular Medicine
The John Radcliffe Hospital
Oxford University
Oxford, UK

Anthony Monaco
Neurogenetics Group
Wellcome Trust Centre for Human Genetics
Oxford, UK

B. Keavney
Institute of Human Genetics
University of Newcastle
Newcastle, UK

Bipen D. Patel
Department of Public Health and Primary Care
Institute of Public Health
Cambridge University
Cambridge, UK

Bruce A.J. Ponder
Cancer Research UK Human Cancer Genetics Group
Department of Oncology
Strangeways Research Laboratory
Cambridge, UK

C. Roland Wolf
CR-UK Molecular Pharmacology Unit
Ninewells Hospital & Medical School
Dundee, UK

Christopher M. Dobson
Department of Chemistry
University of Cambridge
Cambridge, UK

David B. Goldstein
Department of Biology (Galton Lab)
University College London
London, UK

David A. Lomas
Respiratory Medicine Unit
Department of Medicine
University of Cambridge
Cambridge Institute for Medical Research
Cambridge, UK

Dirk-Jan Kleinjan
MRC Human Genetics Unit
Western General Hospital
Edinburgh, UK

Donald F. Conrad
Department of Human Genetics
The University of Chicago
Chicago IL
USA

Donncha Dunican
MRC Human Genetics Unit
Medical Research Council
Western General Hospital
Edingburgh, UK

D.R. Higgs
MRC Molecular Haematology Unit
Weatherall Institute of
Molecular Medicine
University of Oxford
John Radcliffe Hospital
Oxford, UK

D. Timothy Bishop
Cancer Research UK
Clinical Centre
St James University Hospital
University of Leeds
Leeds, UK

Gabrielle Barnby
Neurogenetics Group
Wellcome Trust Centre for Human Genetics
Oxford, UK

Geoffrey Pradella
AHRC Research Centre for Studies in
Intellectual Property
and Technology Law
University of Edinburgh
Edinburgh, UK

Gianpiero L. Cavalleri
Department of Biology (Galton Lab)
University College London
London, UK

Gillian Smith
CR-UK Molecular Pharmacology Unit
Ninewells Hospital & Medical School
Dundee, UK

Harry Campbell
Department of Public Health Sciences
University of Edinburgh
Edinburgh, UK

I. Sadaf Farooqi
CIMR
Wellcome Trust/MRC Building
Addenbrookes' Hospital
Cambridge, UK

Igor Rudan
School of Public Health Andrija Stampar
University of Zagreb
Zagreb, Croatia

Jean-Pierre Hugot
Department of Paediatric
Gastroenterology
INSERM
Hopital Robert Debré
Paris, France

John I. Bell
The Churchill Hospital
University of Oxford
Headington
Oxford, UK

John Trowsdale
Immunology Division
Department of Pathology
Cambridge, UK

Jonathan K. Pritchard
Department of Human Genetics
The University of Chicago
Chicago IL
USA

Jonathan Rees
Department of Dermatology
University of Edinburgh
Edinburgh, UK

Karen P. Steel
Wellcome Trust Sanger Institute
Cambridge, UK

K. B. G. Dear
National Centre for Epidemiology and
Population Health
The Australian National University
Canberra, Australia

Kopal Tandon
Neurogenetics Group
Wellcome Trust Centre for Human Genetics
Oxford, UK

Lars Fugger
The Churchill Hospital
University of Oxford
Headington
Oxford, UK

Malcolm G. Dunlop
MRC Human Genetics Unit
Western General Hospital
Edinburgh, UK

Mark Chamberlain
CR-UK Molecular Pharmacology Unit
Ninewells Hospital & Medical School
Dundee, UK

M. Lathrop
Centre National de Genotypage
France

Mark I. McCarthy
Oxford Centre for Diabetes,
Endocrinology & Metabolism
Churchill Hospital Site
Headington
Oxford, UK

Naomi R. Wray
Queensland Institute of Medical Research
PO Royal Brisbane Hospital
Brisbane, Australia

Nicholas D. Hastie
MRC Human Genetics Unit
Western General Hospital
Edinburgh, UK

Paul Cook
Division of Clinical Sciences
Imperial College
London, UK

Paul D. P. Pharoah
Cancer Research UK Human Cancer
Genetics Group
Department of Oncology
Strangeways Research Laboratory
Worts Causeway
Cambridge, UK

Peter H. St George-Hyslop
Department of Medicine
Division of Neurology
The Toronto Hospital
University of Toronto
Toronto, Canada

Peter McGuffin
MRC Social, Genetic and Developmental
Psychiatry Centre
Institute of Psychiatry
King's College
London, UK

Peter M. Visscher
Queensland Institute of Medical Research
PO Royal Brisbane Hospital
Brisbane, Australia

Pierre Hainaut
International Agency for Research on Cancer
Lyon, France

Renate Gertz
Generation Scotland
AHRC Research Centre for Studies in
Intellectual Property
and Technology Law
University of Edinburgh
Edinburgh, UK

Richard Meehan
MRC Human Genetics Unit
Western General Hospital
Edinburgh, UK

Robert A. Colbert
William S Rowe Division of Rheumatology
Department of Paediatrics
Cincinnati Children's Hospital Medica Center and
The University of Cincinnati
Cincinnati, USA

Rossi Naoumova
Division of Clinical Sciences
Imperial College
London, UK

Sari Pennings
Molecular Physiology
University of Edinburgh
Edinburgh, UK

Shawn Harmon
INNOGEN
ESRC Centre for Social and Economic
Research on Innovation in
Genomics
University of Edinburgh, UK

Sridevi Kalidindi
Neurogenetics Group
Wellcome Trust Centre for Human Genetics
Oxford, UK

Stephen O'Rahilly
CIMR
Wellcome Trust/MRC Building
Addenbrookes' Hospital
Cambridge, UK

Stephen P. Robertson
Department of Paediatrics and
Child Health
Dunedin School of Medicine
Dunedin, New Zealand

Stuart A. Cook
Division of Clinical Sciences
Imperial College
London, UK

Susan M. Farrington
Colon Cancer Genetics Group
Department of Surgery
University of Edinburgh
Edinburgh, UK

Thomas T. Perls
Boston University Medical Center
Boston MA
USA

Timothy J. Aitman
Division of Clinical Sciences
Imperial College
London, UK

W. G. Wood
MRC Molecular Haematology Unit
Weatherall Institute of Molecular Medicine
University of Oxford
John Radcliffe Hospital
Oxford, UK

Foreword

The announcement of the partial completion of the Human Genome Project was accompanied by expansive claims about the impact that this remarkable achievement will have on medical practice in the near future. The media and even some of the scientific community suggested that, within the next 20 years, many of our major killers, at least those of the rich countries, will disappear. What remains of day-to-day clinical practice will be individualized, based on a knowledge of a patient's particular genetic make-up, and survival beyond 100 years will be commonplace. Indeed, the hyperbole continues unabated; as I write a British newspaper announces that, based on the results of manipulating genes in small animals, future generations of humans can look forward to lifespans of 200 years.

This news comes as something of a surprise to the majority of practicing doctors. The older generation had been brought up on the belief that most diseases are environmental in origin and that those that are not, vascular disease and cancer for example, can be lumped together as ''degenerative'', that is the natural consequence of increasing age. More recent generations, who know something about the interactions between the environment and vascular pathology and are aware that cancer is the result of the acquisition of mutations of oncogenes, still believe that environmental risk factors are the major cause of illness; if we run six miles before breakfast, do not smoke, imbibe only homeopathic doses of alcohol, and survive on the same diets as our

hunter-gatherer forebears, we will grow old gracefully and live to a ripe old age. Against this background it is not surprising that today's doctors were astonished to hear that a knowledge of our genetic make-up will transform their practice almost overnight.

The rather exaggerated claims for the benefits of genomics for clinical practice stem from the notion that, since twin studies have shown that there is a variable genetic component to most common diseases, the definition of the different susceptibility genes involved will provide a great deal of information about their pathogenesis and, at the same time, offer the pharmaceutical industry many new targets for their management. An even more exciting prospect is that it may become possible to identify members of the community whose genetic make-up renders them more or less prone to noxious environmental agents, hence allowing public health measures to be focused on subgroups of populations. And if this is not enough, it is also claimed that a knowledge of the relationship between drug metabolism and genetic diversity will revolutionize clinical practice; information about every patient's genome will be available to their family practitioners, who will then be able to adjust the dosage of their drugs in line with their genetic constitution.

Enough was known long before the completion of the Genome Project to suggest that the timescale of this rosy view of genomics and health is based more on hope than reality. For example, it was already clear that the remarkable phenotypic diversity of single gene disorders, that is those whose inheritance follows a straightforward Mendelian pattern, is based on layer upon layer of complexity, reflecting multiple modifier genes and complex interactions with the environment. Even after the fruits of the Genome Project became available, and although there were a few successes, genome-wide searches for the genes involved in modifying an individual's susceptibility to common diseases often gave ambiguous results. Similarly, early hopes that sequence data obtained from pathogen genomes, or those of their vectors, would provide targets for drug or vaccine development have been slow to come to fruition. And while there have been a few therapeutic successes in the cancer field – the development of an agent directed at the abnormal product of an oncogene in a common form of human leukemia for example – an increasing understanding of the complexity of neoplastic transformation at the molecular level has emphasized the problems of reversing this process.

In retrospect, none of these apparent setbacks should have surprised us. After all, it seems likely that most common diseases, except monogenic disorders, reflect a complex interplay between multiple and variable environmental factors and the individual responses of patients which are fine-tuned by the action of many different genes, at least some of which may have very small phenotypic effects. Furthermore, many of the refractory illnesses, particularly those of the rich countries, occur in middle or old age and hence the ill-understood biology of aging adds yet another level of complexity to their pathogenesis. Looked at in this way, it was always unlikely that there would be any quick answers to the control of our current killers.

Because the era of molecular medicine is already perceived as a time of unfulfilled promises, in no small part because of the hype with which it was heralded, the field is being viewed with a certain amount of scepticism by both the medical world and the community at large. Hence, this book, which takes a hard-headed look at the potential of the role of genetics for the future of medical practice, arrives at a particularly opportune time. The editors have amassed an excellent team of authors, all of whom are leaders in their particular fields and, even more importantly, have worked in them long enough to be able to place their potential medical roles into genuine perspective. Furthermore, by presenting their research in the kind of language which will make their findings available to practising doctors, they have performed an invaluable service by interpreting the complexities of genomic medicine for their clinical colleagues.

The truth is that we are just at the beginning of the exploration of disease at the molecular level and no-one knows where it will lead us in our search for better ways of controlling and treating common illness, either in the developing or developed countries. In effect, the position is very similar to that during the first dawnings of microbiology in the second half of the nineteenth century. In March 1882, Robert Koch announced the discovery of the organism that causes tuberculosis. This news caused enormous excitement throughout the world; an editorial writer of the London *Times* newspaper assured his readers that this discovery would lead immediately to the treatment of tuberculosis, yet 62 frustrating years were to elapse before Selman Waksman's announcement of the development of streptomycin. There is often a long period between major discoveries in the research laboratory and their application in the clinic; genomics is unlikely to be an exception.

Those who read this excellent book, and I hope that there will be many, should be left in no doubt that the genetic approach to medical research and practice offers us the genuine possibility of understanding the mechanisms that underlie many of the common diseases of the richer countries, and, at the same time, provides a completely new approach to attacking the major infectious diseases which are decimating many of the populations of the developing countries. Since we have no way of knowing the extent to which the application of our limited knowledge of the environmental causes of these diseases to their control will be successful, it is vital that we make full use of what genomics has on offer.

We are only witnessing the uncertain beginnings of what is sure to be an extremely exciting phase in the development of the medical sciences; scientists should constantly remind themselves and the general public that this is the case, an approach which is extremely well exemplified by the work of the editors and authors of this fine book. I wish them and their publisher every success in this new venture.

D. J. Weatherall
Oxford

SECTION 1

Introductory principles

1

Genes and their expression

Dirk-Jan Kleinjan

The completion of the human genome project has heralded a new era in biology. Undoubtedly, knowledge of the genetic blueprint will expedite the search for genes responsible for specific medical disorders, simplify the search for mammalian homologues of crucial genes in other biological systems and assist in the prediction of the variety of gene products found in each cell. It can also assist in determining the small but potentially significant genetic variations between individuals. However, sequence information alone is of limited value without a description of the function and, importantly, of the regulation of the gene products. Our bodies consist of hundreds of different cell types, each designed to perform a specific role that contributes to the overall functioning of the organism. Every one of these cells contains the same 20 000 to 30 000 genes that we are estimated to possess. The remarkable diversity in cell specialization is achieved through the tightly controlled expression and regulation of a precise subset of these genes in each cell lineage. Further regulation of these gene products is required in the response of our cells to physiological and environmental cues. Most impressive perhaps is how a tightly controlled program of gene expression guides the development of a fertilised oocyte into a full-grown adult organism. The human genome has been called our genetic blueprint, but it is the process of gene expression that truly brings the genome to life. In this chapter we aim to provide a general overview of the physical appearance of genes and the mechanisms of their expression.

What is a gene?

The realization that certain traits are inherited from our ancestors must have been around for ages, but the study of these hereditary traits was first established by the Austrian monk Gregor Mendel. In his monastery in Brno, Czechoslovakia, he performed his famous experiments crossing pea plants and following a number of hereditary traits. He realised that many of these traits were under the control of two distinct factors, one coming from the male parent and one from the female. He also noted that the traits he studied were not linked and thus must have resided on separate hereditary units, now known as chromosomes, and that some appearances of a trait could be dominant over others. In the early 1900s, with the rediscovery of Mendel's work, the factors conveying hereditary traits were named ''genes'' by Wilhelm Johanssen. A large amount of research since then has led to our current understanding about what constitutes a gene and how genes work.

Genes can be defined in two different ways: the gene as a ''unit of inheritance'', or the gene as a physical entity with a fixed position on the chromosome that can be mapped in relation to other genes (the genomic locus). While the latter is the more traditional view of a gene the former view is more suited to our current understanding of the genomic architecture of genes. A gene gives rise to a phenotype through its ability to generate an RNA (ribonucleic acid) or protein product. Thus the

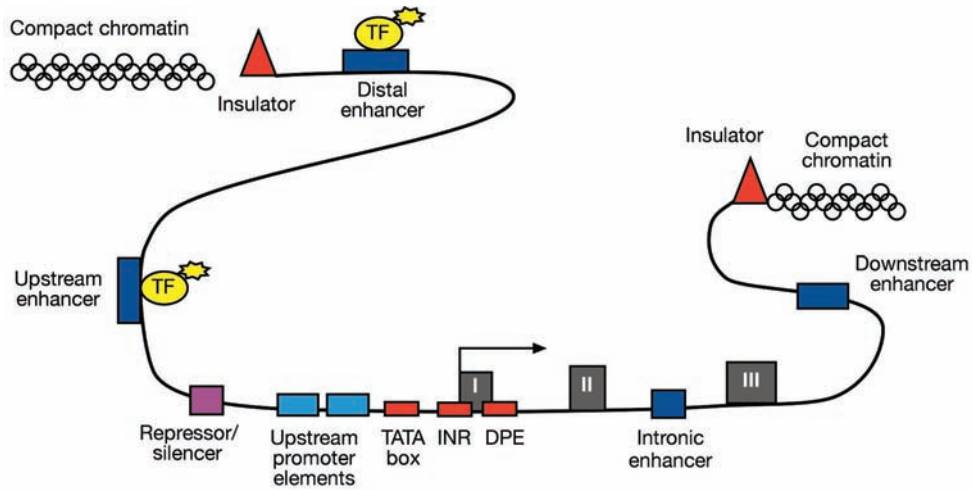

Figure 1.1 The chromosomal architecture of a (fictional) eukaryotic gene. Depicted here is a gene with three exons (grey boxes with roman numerals) flanked by a complex arrangement of *cis*-regulatory elements. The functions of the various elements are explained in the text.

functional genetic unit must encompass not only the DNA (deoxyribonucleic acid) that is transcribed into RNA, but also all of the surrounding DNA sequences that are involved in its transcription. Those regulatory sequences are called the *cis*-regulatory elements, and contain the binding sites for *trans*-acting transcription factors. *Cis*-regulatory elements can be grouped into different classes which will be discussed in more detail later. Recently it has become recognized that *cis*-regulatory elements can be located anywhere on the chromosomal segment surrounding the gene from next to the promoter to many hundreds of kilobases away, either upstream or downstream. Notably, they can also be found in introns of neighboring genes or in the intergenic region beyond the next gene. Crucially, the concept of a gene as a functional genetic unit allows genes to overlap physically yet remain isolated from one another if they bind different sets of transcription factors (Dillon, 2003). As more genes are characterized in greater detail, it is becoming clear that overlap of functional genetic units is a widespread phenomenon.

The transcriptome and the proteome

An enormous amount of knowledge has been gained about genes since they were first discovered, including the fact that at the DNA level most genes in eukaryotes are split, i.e. they contain exons and introns (Berget *et al.*, 1977; Chow *et al.*, 1977) (Figure 1.1). The introns are removed from the RNA intermediate during gene expression in a process called RNA splicing. The split nature of many genes allows the opportunity to create multiple different messages through various mechanisms collectively termed alternative splicing (Figure 1.2). A fully detailed image of a complex organism requires knowledge of all the proteins and RNAs produced from its genome. This is the goal of proteomics, the study of the complete protein sets of all organisms. Due to the existence of alternative splicing and alternative promoter usage in many genes the complement of RNAs and proteins of an organism far exceeds the total number of genes present in the genome. It has been estimated that at least 35% of all human genes show variably spliced products (Croft *et al.*, 2000). It is not uncommon to see genes

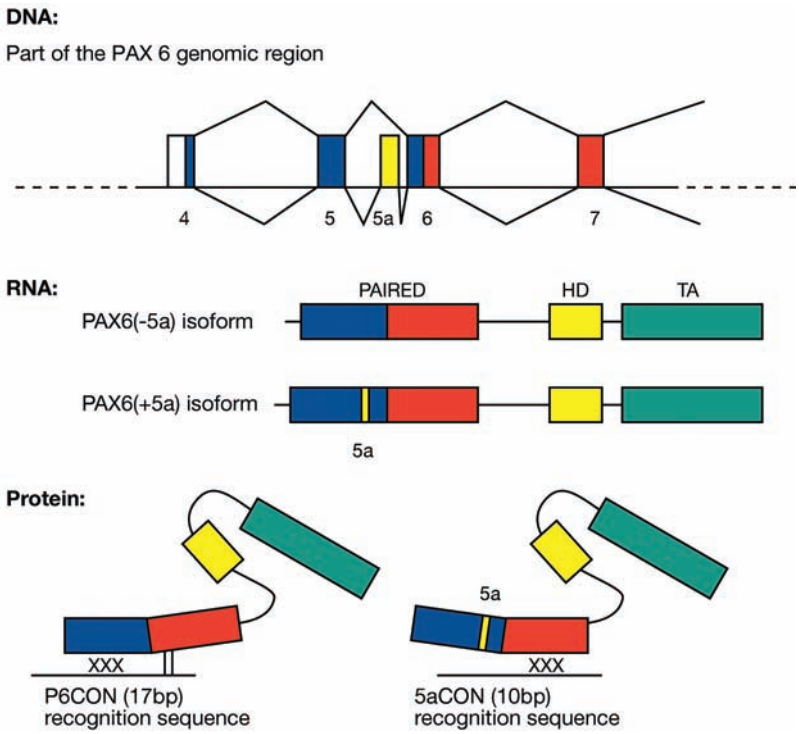

Figure 1.2 The impact of alternative splicing. As an example, part of the genomic region of the *PAX6* transcription factor gene, which has an alternative exon 5a, is shown. The inclusion or exclusion of this exon in the mRNA generates two distinct isoforms, PAX6(+5a) and PAX6(−5a). As a result of the inclusion of exon 5a an extra 14 amino acids are inserted into the paired box (PAIRED), one of its two DNA binding domains, the other being the homeobox domain (HD). The transactivation domain (TA) is also shown. This changes the conformation of the paired box causing it to bind to a different recognition sequence (5aCON) that is found in a different subset of target genes, compared with the –5a isoform recognition sequence (P6CON) (Epstein *et al.*, 1994).

with a dozen or more different transcripts. There are also remarkable examples of hundreds or even thousands of functionally divergent mRNAs (messenger RNAs) and proteins being produced from a single gene. In the human genome such transcript-rich genes include the neurexins, N-cadherins and calcium-activated potassium channels (e.g. Rowen *et al.*, 2002). Thus the estimated 35 000 genes in the human genome could easily produce several hundred thousand proteins or more.

Variation in mRNA structure can be brought about in many different ways. Certain exons can be spliced in or skipped. Introns that are normally excised can be retained in the mRNA. Alternative 5′ or 3′ splice sites can be used to make exons shorter or longer. In addition to these changes in splicing, use of alternative promoters (and thus start sites) or alternative polyadenylation sites also allows production of multiple transcripts from the same gene. (Smith and Valcarcel, 2000). The effect which these alternative splice events can have on the structure of the resulting protein is similarly diverse. Functional domains can be added or left out of the encoded protein. Introduction of an early stop codon can result in a truncated protein or an unstable RNA. Short peptide sequences can be included in the protein that can have very specific

effects on the activity of the protein, e.g. they can change the binding specificity of transcription factors or the ligand binding of growth factor receptors. The inclusion of alternative exons can lead to a change in the subcellular localization, the phosphorylation potential or the ability to form protein–protein interactions. The *DSCAM* gene of *Drosophila* provides a particularly striking example of the number of proteins that can be generated from a single gene. This gene, isolated as an axon guidance receptor responsible for directing axon growth cones to their targets in the Bolwig organ of the fly, has 24 exons. However, 4 of these exons are encoded by arrays of potential alternative exons, used in a mutually exclusive manner, with exon 4 having 12 alternatives, exon 6 having 48 alternatives, exon 9 having 33 alternatives and exon 17 having another 2. Thus, if all of the possible combinations were used, the *DSCAM* gene would produce 38 016 different proteins (Schmucker *et al.*, 2000). This is obviously an extreme example, but it highlights the fact that gene number is not a reliable marker of the protein complexity of an organism. Additional functional variation comes from post-translational modification. Post-translational modifications are covalent processing events which change the properties of a protein by proteolytic cleavage or by addition of a modifying group to one or more amino acids (e.g. phosphorylation, glycosylation, acetylation, acylation and methylation). Far from being mere ''decorations,'' post-translational modification of a protein can finely tune the cellular functions of each protein and determine its activity state, localization, turnover, and interactions with other proteins.

Gene expression

The first definition of the gene as a functional unit followed from the discovery that individual genes are responsible for the production of specific proteins. The difference in chemical nature between the DNA of the gene and its protein product led to the concept that a gene codes for a protein. This in turn led to the discovery of the complex apparatus that allows the DNA sequence of a gene to generate an RNA intermediate which in turn is processed into the amino acid sequence of a protein. This sequence of events from DNA to RNA to protein has become known as the central dogma of molecular biology. Recent progress has revealed that many of the steps in the pathway from gene sequence to active protein are connected. To provide a framework for the large number of events required to generate a protein product we will follow a generalized pathway from gene to protein as follows.

The gene expression pathway usually starts with an initial signal, e.g. cell cycle progression, differentiation, hormonal stimulation. The signal is conveyed to the nucleus and leads to activation of specific transcription factors. These in turn bind to *cis*-regulatory elements, and, through interaction with other elements of the transcription machinery, promote access to the DNA (chromatin remodelling) and facilitate the recruitment of the RNA polymerase enzymes to the transcription initiation site at the core promoter. In eukaryotes there are three RNA polymerases (RNAPs; see also below). Here we will focus on the expression of genes transcribed by RNAPII, although many of the same basic principles apply to the other polymerases. Soon after RNAP II initiates transcription, the nascent RNA is modified at its 5′ end by the addition of a ''cap'' structure. This ^{7Me}G cap serves to protect the new RNA transcript from attack by nucleases and later acts as a binding site for proteins involved in nuclear export to the cytoplasm and in its translation (Proudfoot, 1997). After the ''initiation'' stage RNAP II starts to move 5′ to 3′ along the gene sequence to extend the RNA transcript in a process called ''elongation''. The elongation phase of transcription is subject to regulation by a family of elongation factors (Uptain *et al.*, 1997). The coding sequences (exons) of most genes are interrupted by long non-coding sequences (introns), which are removed by the process of mRNA splicing. When RNAP II reaches the end of a gene it stops transcribing

(''termination''), the newly synthesized RNA is cleaved off (''cleavage'') and a polyadenosine tail is added to the 3′ end of the transcript ('polyadenylation') (Proudfoot, 1997).

As transcription occurs in the nucleus and translation in the cytoplasm (though some sort of translation proofreading is thought to occur in the nucleus, as part of the ''nonsense-mediated decay'' process, see below), the next phase is the transport of the transcript to the cytoplasm through pores in the nuclear membrane. This process is mediated by factors that bind the mRNA in the nucleus and direct it into the cytoplasm through interaction with proteins that line the nuclear pores (Reed and Hurt, 2002). Translation of mRNA takes place on large ribonucleoprotein complexes called ribosomes. It starts with the localization of the start codon by translation initiation factors and subunits of the ribosome and once again involves elongation and termination phases (Dever, 2002). Finally the nascent polypeptide chain undergoes folding, in some cases assisted by chaperone proteins, and often post-translational modification to generate the active protein.

The process of nonsense-mediated mRNA decay (NMD) is increasingly recognized as an important eukaryotic mRNA surveillance mechanism that detects and degrades mRNAs with premature termination codons (PTC+ mRNAs), thus preempting translation of potentially dominant-negative, carboxyl-terminal truncated proteins (Maquat, 2004). It has been known for more than a decade that nonsense and frameshift mutations which induce premature termination codons can destabilize mRNA transcripts in vivo. In mammals, a termination codon is recognized as premature if it lies more than about 50 nucleotides upstream of the final intron position, triggering a series of interactions that leads to the decapping and degradation of the mRNA. Although still controversial, it has been suggested that for some genes regulated alternative splicing is used to generate PTC+ mRNA isoforms as a means to downregulate protein expression, as these alternative mRNA isoforms are degraded by NMD rather than translated to make protein. This system has been termed regulated unproductive splicing and translation (RUST) (Neu-Yilik *et al.*, 2004; Sureau *et al.*, 2001; Lamba *et al.*, 2003).

Transcriptional regulation

As follows clearly from the previous section, the expression of a gene can be regulated at several stages in the process from DNA to protein product: at the level of transcription; RNA stability and export; and at the level of translation or post-translational modification or folding. However, for most genes transcriptional regulation is the main stage at which control of expression takes place. In this section we take a more detailed look at the issues involved in RNAPII transcription.

Promoters and the general transcription machinery

Gene expression is activated when transcription factors bind to their cognate recognition motifs in gene promoters, in interaction with factors bound at *cis*-regulatory sequences such as enhancers, to form a complex that recruits the transcription machinery to a gene. A typical core promoter encompasses 50–100 basepairs surrounding the transcription start site and forms the site where the pre-initiation complex, containing RNAPII, the general transcription factors (GTFs) and coactivators, assemble. The promoter thus positions the start site as well as the direction of transcription. The core promoter alone is generally inactive in vivo, although it may support low or basal levels of transcription in vitro. Activators greatly stimulate transcription levels and the effect is called activated transcription.

The pre-initiation complex that assembles at the core promoter consists of two classes of factors: (1) the GTFs including RNAPII, TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH (Orphanides *et al.*, 1996) and (2) the coactivators and corepressors

that mediate the response to regulatory signals (Myer and Young, 1998). In mammalian cells those coactivator complexes are heterogeneous and sometimes purify as a separate entity or as part of a larger RNAPII holoenzyme. The first step in the assembly of the pre-initiation complex at the promoter is the recognition and binding of the promoter by TFIID. TFIID is a multisubunit protein containing the TATA binding protein (TBP) and 10 or more TBP-associated factors (TAF_{II}s). A number of sequence motifs have been identified that are typically found in core promoters and are the recognition sites for TFIID binding: (1) the TATA box, usually found 25–30 BP upstream of the transcription start site and recognized by TBP, (2) the initiator element, (INR) overlapping the start site, (3) the downstream promoter element or DPE, located approximately 30 BP downstream of the start, (4) the TFIIB recognition element, found just upstream of the TATA box in a number of promoters (Figure 1.1). Most transcriptionally regulated genes have at least one of the above motifs in their promoter(s). However, a separate class of promoter, which is often associated with ubiquitously expressed ''housekeeping genes'', appears to lack these motifs but instead is characterized by a high G/C content and multiple binding sites for the ubiquitous transcription factor Sp1 (Smale, 2001; Smale and Kadonaga, 2003).

RNA polymerases

In eukaryotes nuclear transcription is carried out by three RNA polymerases, I, II and III, which can be distinguished by their subunit composition, drug sensitivity and nuclear localization. Each polymerase is specific to a particular class of target genes. RNAP I is localized in the nucleoli, where multiple enzymes simultaneously transcribe each of the many active 45S rRNA genes required to maintain ribosome numbers as cells proliferate. RNAPs II and III are both localized in the nucleoplasm. RNAP II is responsible for the transcription of protein-encoding mRNA as well as snRNAs and a growing number of other non-coding RNAs. RNAP III transcribes genes encoding other small structural RNAs, including tRNAs and 5S RNA. Each of the polymerases has its own set of associated GTFs.

RNAP II is an evolutionarily conserved protein composed of two major, specific subunits, RPB1 and RPB2, in conjunction with 10 smaller subunits. RPB1 contains an unusual carboxy-terminal domain (CTD), composed in mammals of 52 repeats of a heptapeptide sequence. Cycles of phosphorylation and dephosphorylation of the CTD play a pivotal role in mediating its function as a nucleating center for factors required for transcription as well as cotranscriptional events such as RNA capping, splicing and polyadenylation. Elongating RNAP II is phosphorylated at the Ser2 residues of the CTD repeats.

The manner in which the transcription machinery is assembled at the core promoter remains somewhat unclear. Initial observations seemed to suggest a stepwise assembly of the various factors at the promoter, starting with binding of TFIID to the TATA box. However, more recent research has focussed on recruitment of a single large complex called the holoenzyme. The latter view would certainly simplify matters, as the holoenzyme provides a single target through which activators bound to an enhancer or promoter can recruit the general transcription machinery (Myer and Young, 1998).

Cis-regulatory elements

Gene expression is controlled through promoter sequences located immediately upstream of the transcriptional start site of a gene, in interaction with additional regulatory DNA sequences that can be found around or within the gene itself. The sequences located in the region immediately upstream of the core promoter are usually rich in binding sites for a subgroup of ubiquitous, sequence-specific transcription factors including Sp1 and CTF/NF-I (CCAAT binding factor). These immediate upstream sequences are usually termed the regulatory promoter, while sequences found

at a greater distance are called *cis*-regulatory elements. Together with the transcribed regions of genes, the promoters and *cis*-regulatory elements form the working parts of the genome. It has been estimated that around 5% of the human genome is under evolutionary constraint, and hence may be assumed to contribute to the fitness of the organism in some way. However, less than a third of this functional DNA comprises coding regions, while the rest is made up of different classes of regulatory elements such as promoters, enhancers and silencers (which control gene expression) and locus control regions, insulators and matrix attachment regions (which mediate chromatin organization). There is, as yet, no clear understanding of how exactly promoters interact with the various *cis*-regulatory elements.

Enhancers and repressors

Enhancers are stretches of DNA, commonly spanning a few hundred base pairs that are rich in binding sites for transcription factors, and which have a (usually positive) effect on the level of gene transcription. Most enhancers are tissue- or cell-type specific: in cells with sufficient levels of cognate binding factors *cis*-regulatory elements are often exposed as sites that are hypersensitive to DNaseI digestion. This supposedly reflects a local rearrangement in nucleosome positioning and/or local chromatin topology. During differentiation, hypersensitive site formation at promoters and enhancers usually precedes transcription. Transcriptional activators that bind to the *cis*-regulatory elements of a gene are modular proteins with distinct domains, including ones for DNA binding and transcriptional activation (''transactivation''). The DNA binding domain targets the activator to a specific sequence in the enhancer, while the transactivation domain interacts with the general transcription machinery to recruit it to the promoter. Efficient binding of transcription factors to an enhancer often requires cooperative combinatorial interaction with other activators having recognition binding sites nearby in the *cis*-regulatory element. With such a combinatorial system many layers of control can be achieved with a relatively small number of proteins and without the requirement that all genes be expressed in the same way. It also provides the plasticity required by metazoans to respond to developmental and environmental cues, and it effectively integrates many different signaling pathways to provide a complex regulatory network based on a finite number of transcription factors. Nevertheless, setting up and maintaining a tightly controlled program of gene expression requires a big input from our genetic resource, which is reflected in the fact that more than 5% of our genes are predicted to encode transcription factors (Tupler *et al.*, 2001).

Mechanisms of repression are generally less well understood than activation mechanisms, mainly because they are more difficult to study. Repression can occur in several ways: (1) through inactivation of an activator by post-transcriptional modification, dimerization or the blocking of its recognition site, (2) through inhibition of the formation of a pre-initiation complex, (3) mediated through a specific *cis*-regulatory repressor element and its DNA binding protein(s).

Locus control regions

In general, locus control regions (LCRs) share many features with enhancers, in that they coincide with tissue-specific hypersensitive sites, bind typical transcription factors and confer high levels of gene expression on their gene(s). However, LCRs subsume the function of enhancers along with a more dominant ''chromatin opening'' activity, i.e. they modulate transcription by influencing chromatin structure through an extended region in which they induce and maintain an enhanced accessibility to transcription factors. This activity is dominant such that it can override any negative effects from neighbouring regions. The defining characteristic of an LCR is its ability to drive copy-number-dependent, position-independent expression of a linked gene in transgenic assays, even

when the transgene has integrated (randomly) in a region of highly repressive centromeric heterochromatin (Fraser and Grosveld, 1998).

Boundary elements/insulators

Cis-regulatory control regions such as enhancers and LCRs can regulate gene expression over large distances, in some cases several hundreds of kilobases away (Lettice *et al.*, 2002). However, where necessary, mechanisms must have evolved to prevent the unwanted activation of adjacent gene loci. Mechanisms affecting how the genome manages to set up independent expression domains often invoke the use of insulators or boundary elements. These are *cis*-elements that are required at the borders of gene domains and thought to prevent the inappropriate effects of distal enhancers and/or encroaching heterochromatin. Elements that fit this profile have been identified and have been shown to function in assays as transcriptionally neutral DNA elements that can block or insulate the action of enhancers when located in between the enhancer and promoter. Similarly they can also block the influence of negative effects, such as mediated by silencers or by spreading of heterochromatin-like repression when flanking a reporter gene in certain assays. Examples of well-studied insulators are the *Drosophila* gypsy and scs/scs elements, and in vertebrates the IGF2/H19 DMR (differentially methylated region) and HS4 of the chicken β-globin locus (Bell *et al.*, 2001). All vertebrate insulators that have been analyzed so far require the binding of a protein called CTCF for its function.

Matrix attachment sites

Matrix or scaffold attachment sites (MAR/SARs), are DNA sequences isolated as fragments that remain attached to nuclear structures after stringent extraction with high salt or detergent. They are usually A/T rich and are thought to be the sequences where DNA attaches to the nuclear matrix, thus forming the looped structures of the chromosome that were once thought to demarcate separate gene domains. In some cases, MARs have been shown to coincide with transcriptional enhancers and insulators, however, it remains to be established whether this is coincidental or if MARs have a real function in transcriptional regulation (Hart and Laemmli, 1998).

A current view of enhancer action

To explain how regulatory elements relay information to their target promoters through nuclear space, three models have been proposed: looping; tracking; and linking. The looping model predicts that an enhancer/LCR with its bound transcription factors loops through nucleoplasmic space to contact the promoter where it recruits or activates the general transcription machinery. Initial contact is supposed to occur through random collision while affinity between bound proteins will determine the duration of the interactions. In contrast, in the tracking (or scanning) model transcription factors assemble on the DNA at the enhancer and then move along the DNA fiber until they encounter their cognate promoter. At first view this model explains more easily how insulators located between enhancer and promoter can block the influence of enhancers on transcription. In the linking model, transcription factors bind at a distant enhancer, from where the signal is propagated via a growing chain of proteins along the DNA towards the promoter.

Recently, two novel techniques, 3C-technology (Tolhuis *et al.*, 2002) and RNA-TRAP (Carter *et al.*, 2002) have provided some evidence for a looping model in the regulation of the multigeneβ-globin locus. In these studies, based on the relative levels of cross-linking between various sites within the globin locus in erythroid cells, a spatial clustering of the *cis*-regulatory elements (including the active gene promoters, LCR and other DNAse hypersensitive (HS) sites) was found, with the intervening DNA and the inactive genes in the locus looping out. In brain tissue where the β-globin cluster is not

expressed, the DNA appeared to adopt a relatively straight conformation. These observations have led to the proposal of an active chromatin hub (ACH), a 3-D structure created by clustering of the relevant control elements and bound factors to create a nuclear environment amenable to gene expression. The tissue-specific formation of an ACH would create a mini ''transcription factory'', a local high concentration of transcription factors for the promoter to interact with. It remains to be seen whether ACH formation is a general phenomenon, but it is an attractive model that can explain the existence of distinct, autonomously controlled expression patterns from overlapping gene domains (de Laat and Grosveld, 2003).

Transcriptional regulation and chromatin remodeling

Chromatin structure

While DNA binding proteins and their interactions with the basic synthetic machinery drive transcription, it is now clear that the efficiency and the precision of this process are strongly influenced by higher nuclear organization. The DNA in our cells is packaged in a highly organized and compact nucleoprotein structure known as chromatin (Figure 1.3). This enables the very long strands of DNA to be packaged in a compact configuration in the nucleus. The basic organizational unit of chromatin is the nucleosome, which consists of 146 bp of DNA wrapped almost twice around a protein core, the histone octamer, containing two copies each of four histone proteins: H2A, H2B, H3 and H4 (Luger, 2003). Histones are small, positively charged proteins which are very highly conserved among eukaryotes. The structure created by the DNA wrapped around the nucleosomes is known as the 10 nm fibre, also referred to as the so-called ''beads on a string'' structure. The linker histones H1 and H5 can be found on the DNA in between the beads and assist in further compaction to create less well-defined levels of higher order chromatin folding (e.g. 30 nm fibre). In addition to histones, several other abundant proteins are commonly associated with chromatin, including various HMG proteins and HP1 (specifically at heterochromatin). Visually ''compact'' chromatin such as found at the centromeres is called heterochromatin. Silenced genes are thought to adopt a comparable compact and relatively inaccessible chromatin structure. Expressed genes tend to reside in what is called euchromatin, where genes and their control elements are more accessible to transcriptional activators by virtue of an open structure. Many aspects of chromatin structure are based on interactions between nucleosomal histones and DNA, neighbouring nucleosomes and the non-histone chromatin binding proteins. Most of these interactions involve the N-terminal tails of the core histones, which protrude from the compact nucleosome core and are among the most highly conserved sequences in eukaryotes. Post-translation modifications of the N-termini, in particular of histones H3 and H4, modulates their interaction potential and hence influences the folding and functional state of the chromatin fibre. Three types of modification are known to occur on histone tails: acetylation, phosphorylation and methylation (Spotswood and Turner, 2002).

Chromatin modification and transcription

To activate gene expression, transcriptional activator proteins must bind to and decompact repressive chromatin to induce transcription. To do so they frequently require the cooperation of the diverse family of transcriptional coactivator proteins, as mentioned earlier. The role of these coactivator protein complexes was initially obscure until it was found that many of them carry subunits that have one of two activities: (1) histone acetyl transferase (HAT) activity, or (2) adenosine triphosphate (ATP)-dependent chromatin remodeling activity.

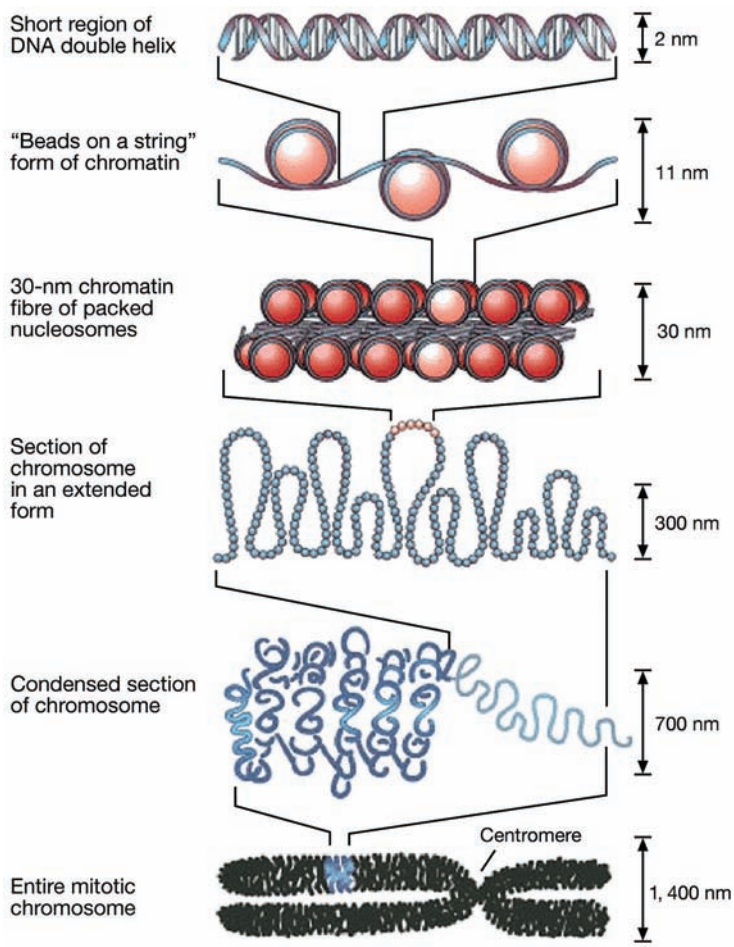

Figure 1.3 DNA packaging. In eukaryotic cells DNA is packaged into a nucleoprotein structure called chromatin. The basic subunit of chromatin is the nucleosome consisting of two superhelical turns of DNA wound around a histone octamer. This 'beads on a string' structure is folded into a 30 nm (diameter) fibre, which is further packaged into so far largely uncharacterised higher-order structures.

Histone acetyl transferase activity

Histone acetylation is an epigenetic mark that is strongly correlated with the transcriptional activity of genes. The acetylation of histones, and its structural effects, can be reversed by the action of dedicated histone deacetylases (HDACs). Thus, the interplay between HATs and HDACs results in dynamic changes in chromatin structure and activity states. Acetylation of lysine residues in the histone tails results in a reduction of their overall positive charge, thus loosening the histone–DNA binding. However, the situation is more complex and the pattern of acetylation at specific lysines appears to be very important. The functional importance of HATs and HDACs is

highlighted by their link with cancer progression (see Chapters 15–17) and their involvement in some human disorders, such as Rubinstein-Taybi and Fragile X syndrome (Timmermann *et al.*, 2001). Acetylation of histones H3 and H4 leads to altered folding of the nucleosomal fibre, thus rendering chromosomal domains more accessible. Consequently, the transcription machinery may be able to access promoters more easily. In addition, the unfolding of chromosomal domains also facilitates the process of transcriptional elongation itself. Nucleosomes form obstacles hindering the progression of RNA polymerase through its template, and the polymerase may need to transfer the nucleosomes to acceptor DNA in the wake of elongation. Thus HATs may also be involved in facilitating the passage of the elongating polymerase, either as part of dedicated elongation factor complexes such as FACT or as an integral activity of the elongation machinery itself (Belotserkovskaya *et al.*, 2004).

Many studies have corroborated the importance of histone acetylation as an epigenetic marker of chromosomal domains. In differentiated higher eukaryotic cells most of the genome exists as hypoacetylated, inactive chromatin. Where this has been studied, activation of housekeeping and cell type-specific genes involves initial acetylation of histones across broad chromatin domains, which is not correlated with active transcription per se, but rather marks a region of transcriptional competence (Bulger *et al.*, 2002). Transcriptional activation within a permissive domain frequently correlates with additional, targeted acetylation of histones at the core promoter (Forsberg and Bresnick, 2001).

Over the past few years histone acetylation has emerged as a central switch between permissive and repressive chromatin structure. More recently other post-transcriptional modifications of residues in the histone tails have also been found to have profound effects on gene transcription, namely ubiquitination, serine phosphorylation, lysine and arginine methylation (Spotswood and Turner, 2002). All these modifications influence each other and rather than just being a means to reorganize chromatin structure they provide a rich source of epigenetic information. The combination of specific histone-tail modifications found on nucleosomes has been suggested to constitute a code that defines the potential or actual transcriptional state. This ''histone code'' is set by specific histone modifying enzymes and requires the existence of non-histone proteins with the capability to read the code (Figure 1.4). The identification of several of the histone modifying enzymes reveals an important further dimension in the control of the structural and functional activities of genes and promoter regulatory elements.

ATP-dependent chromatin remodeling activity

A second important activity carried out by a separate class of transcriptional coactivator complexes is the ATP-dependent remodeling of the chromatin surrounding the promoters of genes, leading to increased mobility and fluidity of local nucleosomes. The prototype of a large family of protein complexes with this function is the SWI/SNF2 complex (other complexes are RSC, CHRAC, NURF, ACF). All these remodeling complexes have at least one subunit with a conserved ATPase motif. The SWI/SNF2 and RSC complexes are thought to use the energy provided by ATP hydrolysis to unwind DNA and displace the nucleosome, while in the case of CHRAC and NURF, sliding of the nucleosomes along the DNA appears to be the mechanism. There is functional interplay between the HAT coactivators and the remodeling complexes, with some evidence from a small number of studies to suggest that histone acetylation precedes SWI/SNF activity and perhaps marks the domain that is to be the substrate for the ATP-dependent remodeling.

More recently a third type of chromatin remodeling has received interest, i.e. the replacement of core histones with non-canonical variants. For example, the histone H3 variant CENP-A is

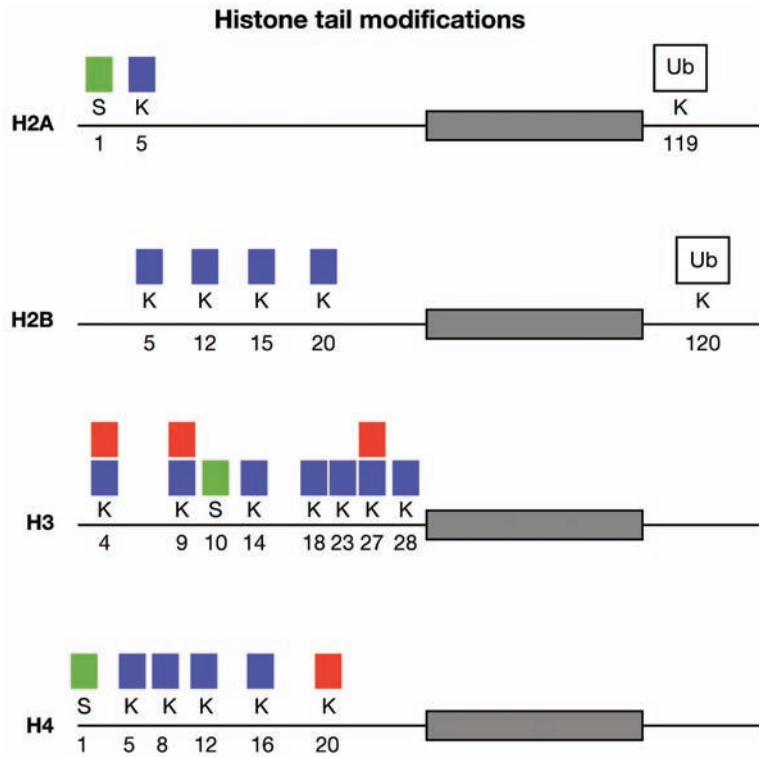

Figure 1.4 The histone code. Schematic representation of the four core histone proteins with their possible modifications. The modifications found on the histones in a particular region of the DNA are thought to provide a code with information on the transcriptional status/competence of the region (the 'histone code'). Some of the modifications shown are mutually exclusive. Blue boxes indicate lysine (K) acetylation, green boxes indicate serine (S) phosphorylation and red boxes indicate lysine methylation. Ub indicates ubiquitination.

associated entirely with centromeric chromatin whereas some evidence suggests that the H3.3 variant replaces H3 in actively transcribed regions (McKittrick *et al.*, 2004). How histone variants substitute for their canonical counterparts remains an intriguing question. Normally nucleosomes are reassembled after cell division, concurrent with the replication process. At actively transcribed regions, nucleosomes are displaced or mobilized to allow access to polymerases and other proteins, and are replaced in a replication-independent process which may have a preference for the incorporation of histone H3.3. A further possibility is the existence of protein complexes that can facilitate the exchange between histones and their variant counterparts in interphase chromatin. Recently a protein, Swr1, a member of the Swi2/Snf2 family, has been identified that can mediate the exchange of histone H2A for its variant H2A.Z in an ATP-dependent fashion in vitro (Kobor *et al.*, 2004).

Epigenetic regulation through DNA modification

Epigenetics is a term used to describe the study of stable alterations to our genetic information that do not involve mutation of the DNA sequence

itself. Two molecular mechanisms are known to mediate epigenetic phenomena: DNA methylation and histone modification (Jaenisch and Bird, 2003) (see Chapter 2). The latter has already been discussed above. DNA methylation in mammals is a post-replication modification that is predominantly found in cytosines of the dinucleotide sequence CpG. Methylation is recognised as a chief contributor to the stable maintenance of silent chromatin. The patterns of DNA methylation are set up and maintained by a family of DNA methyltransferases (DNMTs). Potential explanations for the evolution of DNA methylation invoke its ability to silence transposable elements, its function as a mediator of developmental gene regulation or its function in reducing transcriptional noise (Bird, 2002). In non-embryonic cells, about 80% of CpGs in the genome are methylated. Interrupting this global sea of genomic methylation are the CpG islands, short sequence domains with a high CpG content that generally (with some exceptions) remain unmethylated at all times, regardless of gene expression. Most of these CpG islands are associated with genes and all are thought to contain promoters. How CpG islands remain methylation-free is still an open question.

A general mechanism through which methylation can repress transcription is by interference with the binding of transcription factors to their binding sites. Several factors are known to be blocked from binding to their recognition site when it is methylated, including the boundary element binding protein CTCF (Ohlsson *et al.*, 2001). However, the major mechanism of methylation-mediated repression is through the recruitment of transcriptional repressor complexes by a family of methyl CpG binding proteins. The proteins in this family, which includes MeCP2, MBD1–4 and an unrelated protein called Kaiso, specifically recognize and bind to methylated CpGs through their methyl-binding domain. Both MeCP2 and MBD2 (a subunit of the MeCP1 complex) have been found to interact with corepressor complexes containing HDACs, making a link between DNA methylation and nucleosome modification to induce the silencing of target genes (Jaenisch and Bird, 2003). Two very specific human syndromes have been shown to be caused by mutations in genes linked to DNA methylation: The neurological disorder Rett syndrome, caused by MeCP2 mutation and ICF syndrome caused by DNMT3B mutation (see Chapter 2). The integration of DNA methylation, histone modification and chromatin remodeling is a complex process that depends on the collaboration of many components of the epigenetic machinery. Transitions between different chromatin states are dynamic and depend on a balance between factors that sustain a silent state (such as methylation, HDACs) and those that promote a transcriptionally active state (e.g. HATs) (Figure 1.5).

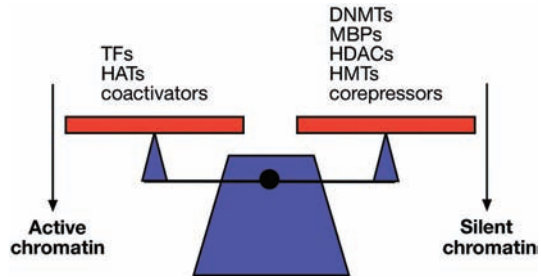

Figure 1.5 The status of the chromatin in a particular region of the genome depends on a balance between factors that sustain a silent state and those that promote a transcriptionally active or competent state. Factors that are correlated with a silent state include DNA methylation by methyltransferases (DNMTs), recognition by methyl binding proteins (MBPs), marking via histone deacetylation by histone deacetylases (HDACs) and histone H3 K9 methylation by histone methyltransferases (HMTs). Transcription factors (TFs), coactivators and histone acetylases (HATs) are involved in promoting an active chromatin state.

Nuclear compartmentalization and dynamics

Further to the influence of epigenetic modifications there is another factor that has been proposed to have an effect on gene expression, and this relates to the positioning of the gene within the nucleus. It has become well accepted that the contents of the nucleus are organized in a highly structured manner. There is emerging recognition that nuclear structure and function are causally inter-related, with mounting evidence for organization of nucleic acids and regulatory proteins into subnuclear domains that are associated with components of nuclear architecture (Spector, 2003).

Mammalian chromosomes occupy discrete regions within interphase nuclei, with little mixing of chromatin between adjacent so-called chromosome ''territories''. Chromosome territories do not occupy specific positions within the nucleus, but a trend has been observed that gene-rich chromosomes tend to be located towards the centre of the nucleus and gene-poor chromosomes towards the periphery (Croft *et al.*, 1999). In some instances, chromatin loops can be seen that appear to escape out of their chromosome territory, and there seems to be a correlation with a high density of active genes on such loops (Williams *et al.*, 2002; Mahy *et al.*, 2002). It remains to be established whether this reflects a specific movement towards an activity providing centre (e.g. a ''transcription factory'') outside the territory. However, there is evidence to suggest that for some genes the activation of gene expression might correspond with a spatial change in gene location from an inactive to an active chromatin compartment in the nucleus (Francastel *et al.*, 2000).

One of the most prominent manifestations of a specialised functional nuclear compartment is the nucleolus, where rRNA synthesis and ribosome biogenesis occurs. Other types of higher order nuclear domains have also been observed including nuclear speckles, interchromatin granule clusters, B-snurposomes, coiled or Cajal bodies and PML bodies (Lamond and Sleeman, 2003). These putative nuclear compartments have been associated with various transcription and RNA processing factors. Discussion is ongoing as to whether these bodies represent active enzymatic centers or inert reservoirs for factors destined for degradation or recycling. One model proposed the existence of transcription factories, localized assemblies of transcription factors and polymerases which draw in nearby active genes requiring to be transcribed (Martin and Pombo, 2003). How these factories stay together is presently unclear and may depend on protein:protein and/or protein:DNA interactions. Whether transcription factories retain their structure in the absence of transcription remains to be seen.

Transcriptional regulation and disease

With such a complex and highly regulated process as transcriptional regulation, the potential for things to go wrong is enormous. Thus a large number of genetic diseases can directly or indirectly be attributed to mutations in components of the gene expression machinery. These vary from mutations in transcription factors and spliceosomal components, to chromatin components and epigenetic factors and finally to mutation or deletion of control elements (Kleinjan and Van Heyningen, 1998; Hendrich and Bickmore, 2001; Gabellini *et al.*, 2003). Furthermore, the potential importance of gene regulation in disease susceptibility and other inherited phenotypes has also become evident in recent years. This has been underlined by the observation that the human genome contains far fewer protein coding genes than expected. Based on this and on the study of some quantitative traits in simpler organisms, it has been proposed that the genetic causes of susceptibility to complex diseases may reflect a different spectrum of sequence variants to the nonsense and missense mutations that dominate

simpler genetic disorders. Amongst this spectrum, polymorphisms that alter gene expression are suspected of playing a prominent role (e.g. Van Laere *et al.*, 2003).

Concluding remarks

The chromosomal domain that contains the information for correct spatial, temporal and quantitative regulation of a particular gene often exceeds the size of the coding region several-fold and may occupy many hundreds of kilobases of DNA. To identify the *cis*-regulatory elements within these large gene domains using classical techniques, such as DNAseI hypersensitive site mapping, footprinting, transfection and in vitro binding assays, is a massive and daunting prospect. For those genes whose function and regulation are conserved in evolution, valuable help is now at hand in the form of comparative genomics. This bioinformatics technique, called ''phylogenetic footprinting'', can be used to identify conserved, non-coding DNA sequences, whose role must subsequently be tested in functional assays. Another factor currently receiving much interest is the role of non-coding RNAs in the mechanisms of gene regulation (Mattick, 2003). However, a eukaryotic gene locus is not just a collection of control elements separated by ''junk'' DNA, but encodes an intricate *cis*-regulatory system consisting of different layers of regulatory information required for the correct output. This information is organized in a defined three-dimensional structure that includes the DNA, chromatin components, and cell-specific as well as general DNA binding and non-DNA binding proteins. The elucidation of the information encoded in these structures and the way it is translated into the enormous complexities of controlled gene expression remains a major challenge for the future.

REFERENCES

Berget, S. M., Moore, C. and Sharp, P. A. (1977). Spliced segments at the 5′ terminus of adenovirus 2 late mRNA. *Proc Natl Acad Sci USA*, **74**, 3171–5.

Bell, A. C., West, A. G. and Felsenfeld, G. (2001). Insulators and boundaries: versatile regulatory elements in the eukaryotic genome. *Science*, **291**, 447–50.

Belotserkovskaya, R., Saunders, A., Lis, J. T. and Reinberg, D. (2004). Transcription through chromatin: understanding a complex FACT. *Biochim Biophys Acta*, **1677**, 87–99.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev*, **16**, 6–21.

Bulger, M., Sawado, T., Schubeler, D. and Groudine, M. (2002). ChIPs of the beta-globin locus: unraveling gene regulation within an active domain. *Curr Opin Genet Dev*, **12**, 170–7.

Carter, D., Chakalova, L., Osborne, C. S., Dai, Y. F. and Fraser, P. (2002). Long-range chromatin regulatory interactions in vivo. *Nat Genet*, **32**, 623–6.

Chow, L. T., Gelinas, R. E., Broker, T. R. and Roberts, R. J. (1977). An amazing sequence arrangement at the 5′ ends of adenovirus 2 messenger RNA. *Cell*, **12**, 1–8.

Croft, J. A., Bridger, J. M., Boyle, S. *et al.* (1999). Differences in the localization and morphology of chromosomes in the human nucleus. *J Cell Biol*, **145**, 1119–31.

Croft, L., Schandorff, S., Clark, F. *et al.* (2000). ISIS, the intron information system, reveals the high frequency of alternative splicing in the human genome. *Nat Genet*, **24**, 340–1.

De Laat, W. and Grosveld, F. (2003). Spatial organization of gene expression: the active chromatin hub. *Chromosome Res*, **11**, 447–59.

Dever, T. E. (2002). Gene-specific regulation by general translation factors. *Cell*, **108**, 545–56.

Dillon, N. (2003). Gene autonomy: positions, please . . . *Nature*, **425**, 457.

Epstein, J. A., Glaser, T., Cai, J. *et al.* (1994). Two independent and interactive DNA-binding subdomains of the Pax6 paired domain are regulated by alternative splicing. *Genes Dev*, **8**, 2022–34.

Forsberg, E. C. and Bresnick, E. H. (2001). Histone acetylation beyond promoters: long-range acetylation patterns in the chromatin world. *Bioessays*, **23**, 820–30.

Francastel, C., Schubeler, D., Martin, D. I. and Groudine, M. (2000). Nuclear compartmentalization and gene activity. *Nat Rev Mol Cell Biol*, **1**, 137–43.

Fraser, P. and Grosveld, F. (1998). Locus control regions, chromatin activation and transcription. *Curr Opin Cell Biol*, **10**, 361–5.

Gabellini, D., Tupler, R. and Green, M.R. (2003). Transcriptional derepression as a cause of genetic diseases. *Curr Opin Genet Dev*, **13**, 239–45.

Hart, C.M. and Laemmli, U.K. (1998). Facilitation of chromatin dynamics by SARs. *Curr Opin Genet Dev*, **8**, 519–25.

Hendrich, B. and Bickmore, W. (2001). Human diseases with underlying defects in chromatin structure and modification. *Hum Mol Genet*, **10**, 2233–42.

Jaenisch, R. and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet*, **33** Suppl, 245–54.

Kleinjan, D.J. and van Heyningen, V. (1998). Position effect in human genetic disease. *Hum Mol Genet*, **7**, 1611–18.

Kobor, M.S., Venkatasubrahmanyam, S., Meneghini, M.D. *et al.* (2004). A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. *PLoS Biol*, **2**, E131.

Lamba, J.K., Adachi, M., Sun, D. *et al.* (2003). Nonsense mediated decay downregulates conserved alternatively spliced ABCC4 transcripts bearing nonsense codons. *Hum Mol Genet*, **12**, 99–109.

Lamond, A.I. and Sleeman, J.E. (2003). Nuclear substructure and dynamics. *Curr Biol*, **13**, R825–8.

Lettice, L.A., Horikoshi, T., Heaney, S.J. *et al.* (2002). Disruption of a long-range cis-acting regulator for Shh causes preaxial polydactyly. *Proc Natl Acad Sci USA*, **99**, 7548–53.

Luger, K. (2003). Structure and dynamic behavior of nucleosomes. *Curr Opin Genet Dev*, **13**, 127–35.

Mahy, N.L., Perry, P.E. and Bickmore, W.A. (2002). Gene density and transcription influence the localization of chromatin outside of chromosome territories detectable by FISH. *J Cell Biol*, **159**, 753–63.

Maquat, L.E. (2004). Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nat Rev Mol Cell Biol*, **5**, 89–99.

Martin, S. and Pombo, A. (2003). Transcription factories: quantitative studies of nanostructures in the mammalian nucleus. *Chromosome Res*, **11**, 461–70.

Mattick, J.S. (2003). Challenging the dogma: the hidden layer of non-protein-coding RNAs in complex organisms. *Bioessays*, **25**, 930–9.

McKittrick, E., Gafken, P.R., Ahmad, K. and Henikoff, S. (2004). Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proc Natl Acad Sci USA*, **101**, 1525–30.

Myer, V.E. and Young, R.A. (1998). RNA polymerase II holoenzymes and subcomplexes. *J Biol Chem*, **273**, 27757–60.

Neu-Yilik, G., Gehring, N.H., Hentze, M.W. and Kulozik, A.E. (2004). Nonsense-mediated mRNA decay: from vacuum cleaner to Swiss army knife. *Genome Biol*, **5**, 218.

Ohlsson, R., Renkawitz, R. and Lobanenkov, V. (2001). CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends Genet*, **17**, 520–7.

Orphanides, G., Lagrange, T. and Reinberg, D. (1996). The general transcription factors of RNA polymerase II. *Genes Dev*, **10**, 2657–83.

Orphanides, G. and Reinberg, D. (2002). A unified theory of gene expression. *Cell*, **108**, 439–51.

Proudfoot, N.J. (1997). 20 years of making messenger RNA. *Trends Genet*, **13**, 430.

Reed, R. and Hurt, E. (2002). A conserved mRNA export machinery coupled to pre-mRNA splicing. *Cell*, **108**, 523–31.

Rowen, L., Young, J., Birditt, B. *et al.* (2002). Analysis of the human neurexin genes: alternative splicing and the generation of protein diversity. *Genomics*, **79**, 587–97.

Schmucker, D., Clemens, J.C., Shu, H. *et al.* (2000). *Drosophila* Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. *Cell*, **101**, 671–84.

Smale, S.T. (2001). Core promoters: active contributors to combinatorial gene regulation. *Genes Dev*, **15**, 2503–8.

Smale, S.T. and Kadonaga, J.T. (2003). The RNA polymerase II core promoter. *Annu Rev Biochem*, **72**, 449–79.

Smith, C.W. and Valcarcel, J. (2000). Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends Biochem Sci*, **25**, 381–8.

Spector, D.L. (2003). The dynamics of chromosome organization and gene regulation. *Annu Rev Biochem*, **72**, 573–608.

Spotswood, H.T. and Turner, B.M. (2002). An increasingly complex code. *J Clin Invest*, **110**, 577–82.

Sureau, A., Gattoni, R., Dooghe, Y., Stevenin, J. and Soret, J. (2001). SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs. *EMBO J*, **20**, 1785–96.

Timmermann, S., Lehrmann, H., Polesskaya, A. and Harel-Bellan, A. (2001). Histone acetylation and disease. *Cell Mol Life Sci*, **58**, 728–36.

Tolhuis, B., Palstra, R. J., Splinter, E., Grosveld, F. and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell*, **10**, 1453–65.

Tupler, R., Perini, G. and Green, M. R. (2001). Expressing the human genome. *Nature*, **409**, 832–3.

Uptain, S. M., Kane, C. M. and Chamberlin, M. J. (1997). Basic mechanisms of transcript elongation and its regulation. *Annu Rev Biochem*, **66**, 117–72.

Van Laere, A. S., Nguyen, M. *et al.* (2003). A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature*, **425**, 832–6.

Williams, R. R., Broad, S., Sheer, D. and Ragoussis, J. (2002). Subchromosomal positioning of the epidermal differentiation complex (EDC) in keratinocyte and lymphoblast interphase nuclei. *Exp Cell Res*, **272**, 163–75.

2

Epigenetic modification of chromatin

Donncha Dunican, Sari Pennings and Richard Meehan

The coding capacity of the human genome is smaller than originally expected; it is predicted that we have 25 000–40 000 genes, only twofold more than a simple organism such as the roundworm *C. elegans* (Pennisi, 2003). This modest increase in gene numbers is counterbalanced by enormous gains in the potential for complex interactions through alternative splicing, and in the regulatory intricacy of elements within and between genes in chromatin (Bentley, 2004) (Chapter 1). Added to this complexity is an increasing repertoire of epigenetic mechanisms which form the basis of gene silencing and genomic imprinting, including DNA methylation, histone modification and RNA interference (RNAi). These mechanisms have profound influences on developmental gene expression and, when perturbed, cancer progression and human disease (Bjornsson *et al.*, 2004; Meehan, 2003).

Location, location, location!

The position of a gene within a eukaryotic chromosome can be a major determinant of its transcriptional properties. In the last century it was shown that the relocation of the white gene from a euchromatic position to a heterochromatic region resulted in its variegated expression in the eye of the fruit fly (*Drosophila melanogaster*) (Dillon and Festenstein, 2002). This observation was an example of epigenetics, which has two closely related meanings: (1) the study of the processes involved in the unfolding development of an organism, including phenomena such as X chromosome inactivation in mammalian females, and the patterning of gene silencing; (2) any mitotically and/or meiotically heritable change in gene function that cannot be explained by changes in DNA sequence (Meehan, 2003; Waddington, 1957). Unlike heterochromatin, which is maintained in a compact and condensed structure throughout the cell cycle, euchromatin undergoes decondensation which is thought to facilitate gene expression. This basic observation of chromatin organization underlies all aspects of epigenetics from molecular biology, molecular cytology and development to clinical medicine (Huang *et al.*, 2003). Model systems in plants, animals and fungi have identified, by genetic and biochemical methods, the dynamic components that facilitate the formation of different types of chromatin, uncovering an increasing array of molecular markers which act as molecular and cytological signatures of either active or inactive chromatin. A major goal is to understand how these different chromatin states are maintained and transmitted, for example at centromeric heterochromatin (Richards and Elgin, 2002), as well as how chromatin structure can transit between active and inactive states. The human diseases that result from mutations in chromatin modifier genes underscore the importance of these molecular processes in normal development. The scope of this review is to give a short introduction to chromatin and illustrate its importance in disease by describing a number of disorders whose

pathology is determined by mutation in genes that are important in chromatin organization. There have been many recent reviews on chromatin-based gene silencing and activating mechanisms (Feinberg *et al.*, 2002; Huang *et al.*, 2003; Jenuwein and Allis, 2001; Lachner *et al.*, 2003; Meehan, 2003; Richards and Elgin, 2002).

Chromatin

The basic repeating unit of chromatin is the nucleosome, which consists of approximately 146 bp of DNA wrapped around an octamer of lysine rich histones (two copies each of histones H2A, H2B, H3 and H4). In metazoans, histone H1 can bind to the DNA in the linker region and contribute to the higher order folding of chromatin (Wolffe, 1998). The basic function of chromatin is to participate in the reversible compaction of DNA (up to 2 metres in length) in the cell into the small nuclear volume (10 microns in diameter) in such a way as to organize and to regulate nuclear processes such as transcription, replication, DNA repair and chromosome segregation. This is achieved by packing the DNA together with histones and non-histone proteins into a series of higher order structures that are dependent on DNA and histone modifications provided by enzymatic remodeling machines. Biochemical fractionation of chromatin into its active and inactive constituents emphasizes that they have different properties (Figure 2.1). In contrast to active euchromatin, inactive heterochromatin is late replicating, has less nuclease accessibility, has more regular nucleosome arrays, contains hypoacetylated histones, histone H3 methylated at lysine 9 (K9) and DNA methylated at 5 methyl-cytosine (m^5C) (Dillon and Festenstein, 2002).

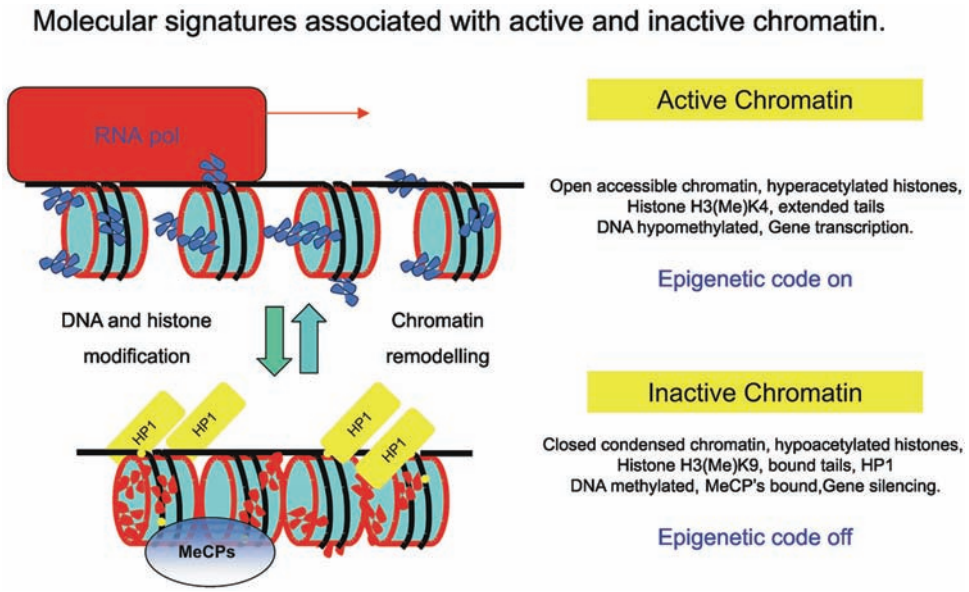

Figure 2.1 Simple model for active versus inactive chromatin. Active chromatin is depicted as being more relaxed (open) with hyperacetylated histone tails (dark blue), facilitating transcription by RNA polymerases. In contrast closed (hetero-) chromatin is not acetylated but instead is associated with lysine 9 methylation on histone H3 which can act as a ligand for histone, DNA methyltransferases and HP1, reinforcing the silent state. This promotes the formation of compact chromatin that is refractory to the transcription complex.

The non-random positioning of nucleosomes over a gene promoter (see Chapter 1) can strategically inhibit transcription initiation and can also ultimately have a bearing on the formation of higher order chromatin structures (Chambeyron and Bickmore, 2004; Gilbert *et al.*, 2004). For example, high affinity nucleosome positioning sites that occlude transcription factor binding sites and may encode the periodicity of regular spaced chromatin (Davey *et al.*, 1995) have been identified on the chicken β-globin promoter. Nucleosome formation and positioning can further be influenced by DNA methylation (Pennings *et al.*, 2005). The discovery that nucleosomes can also be mobile, promotes a dynamic view of chromatin organization based on the regulated positioning of nucleosomes on DNA (Meersseman *et al.*, 1992). Depletion experiments in *Saccharomyces cerevisiae* have demonstrated that loss of histone H4 protein results in increased expression of 15% and decreased expression of 10% of the yeast genes, indicating that histones can have a gene-specific rather than a general repressor role in this organism (Wyrick *et al.*, 1999).

Histone modification

The core histones have long N-terminal regions (tails) extending outward from the nucleosome that can bind to linker DNA and stabilize higher-order oligomeric structures (Kornberg and Lorch, 1999; Luger *et al.*, 1997; Luger and Richmond, 1998). Enzymatic modification of the histone tails can alter their ability to interact both with DNA and with nuclear factors such as Heterochromatin Protein 1 (HP1) (Maison and Almouzni, 2004). There is a general correlation between acetylation of the N-terminal tails and a more open chromatin structure that facilitates gene expression (Figure 2.1). Acetylated nucleosomes have a reduced affinity for DNA resulting in chromatin decompaction, and the acetylation state of lysine residues is dynamically organized via an interplay of histone acetylases (HATs) and deacetylases (HDACs) (Turner, 2000). The targeting of these activities to specific chromosomal loci by sometimes shared transcription factors is a determinant of gene activity and chromatin structure. A variety of HATs has been identified in different nuclear processes with different substrate specificities (Table 2.1). Treatment of pharmacological inhibitors of HDACs can lead to hyper-acetylation of histones, activation of gene expression and decondensation of chromatin in certain test systems (Chambeyron and Bickmore, 2004; Maison and Almouzni, 2004). It is probable that aberrant targeting of chromatin modifying complexes plays a role in the molecular etiology of many diseases, including cancer. For example, the retinoblastoma protein (pRB) recruits HDACs to the transcription factor E2F-driven promoters during the G1 phase of the cell cycle. In wild type cells, E2F controls the expression of a group of cell cycle checkpoint genes whose products are required either for the G1-to-S transition itself or for DNA replication. Inactivation of the retinoblastoma (*Rb*) gene results in loss of silencing of these genes and contributes

Table 2.1. Substrate specificity of lysine histone acetyltransferases (HAT)

Acetyltransferase	Specificity	Consequence
Taf II250	Histone H3: K14	Transcription activation
p300	Histone H3: K14	Transcription activation
PCAF	Histone H3: K14	Transcription activation
p300	Histone H3: K18	Transcription activation
HAT1	Histone H4: K5	Histone deposition
ATF2	Histone H4: K8	Sequence specific Transcription regulation
ATF2	Histone H4: K16	Sequence specific Transcription regulation

Based on Lachner, O'Sullivan and Jenuwein, 2003.

Table 2.2. Substrate specificity of lysine histone methyltransferases (HMT)

Methyltransferase	Specificity	Transcription
Set 7/9	Histone H3: K4	Activation
SUV39H1	Histone H3: K9	Repression
EZH2	Histone H3: K27	Repression
Pr-SET7	Histone H4: K20	Repression
Eu-HMTase1	Histone H3: K9	Repression
SETDB1	Histone H3: K9	Repression

Based on Lachner, O'Sullivan and Jenuwein, 2003.

to an increased proliferative potential and absence of cell cycle checkpoints both in retinoblastoma and many cancer cells (Johnston *et al.*, 2003).

A variety of additional modifications that occur on histones has been identified by genetic and biochemical means, including methylation, phosphorylation, ribosylation, biotinylation and ubiquination (Jenuwein and Allis, 2001; Rodriguez-Melendez and Zempleni, 2003). On this basis, a histone code has been postulated, which suggests that each modification (or a combination of them) has a functional effect, on transcription and/or chromatin organization (Khorasanizadeh, 2004). The code hypothesis also invokes regulatory proteins (modifying enzymes) and effector molecules that interpret the modification pattern present on chromatin such that inactive or active chromatin regions can be distinguished and maintained within or as specific nuclear compartments.

In *Drosophila*, a collection of mutations have been isolated that either enhance or suppress position effect variegation (PEV) of the white-eye reporter gene that is located near heterochromatic sites. *Su(var)2–5*, a suppressor mutation that encodes HP1, has been shown to localize to centromeric heterochromatin in a variety of species (Maison and Almouzni, 2004). Genetic experiments have established that the centromeric localization of HP1 in *Drosophila* is dependent on another suppressor *Su(var)3–9*, which was shown to encode a histone methyltransferase (HMT) that selectively methylates K9 on histone H3 via its SET domain (Rea *et al.*, 2000). A number of lysine-specific HMTs have been identified in humans (Table 2.2), many of which target K9 on histone H3 resulting in mono- di- or tri-methylation (Zhang *et al.*, 2003) with different functional consequences (Lehnertz *et al.*, 2003; Mermoud *et al.*, 2002; Nielsen *et al.*, 2001). K9 methylation creates a low affinity ligand (K_D 10^{-6} M) on H3 for the HP1 protein that binds through a highly conserved protein domain (chromodomain). Under physiological conditions however, HP1 cannot bind oligonucleosomes in vitro even though they contain histone H3 that is di- and tri-methylated at K9 (Meehan *et al.*, 2003b). This is probably due to the very high affinity of N-terminal histone tails for linker DNA. The targeting of HP1 to methylated K9 of H3 is not absolute; instead this interaction might occur de novo during chromatin replication and guide HP1 to heterochromatic regions, prior to the association of H3 tails with the linker DNA (Cowell *et al.*, 2002; Gilbert *et al.*, 2003; Meehan *et al.*, 2003; Quivy *et al.*, 2004). HP1 contains potent protein–protein interaction domains and can be involved in targeting of histone and DNA methyltransferase activities, in addition to nuclear factors such as MBD1 and MeCP2. Different classes of HMT have been identified that modify lysine or arginine, which can stimulate or repress gene expression in different chromatin contexts (Marmorstein, 2003). For example *DOT1* in *Saccharomyces cerevisiae* methylates lysine 79 (MeK79) on histone H3 in bulk chromatin but not at telomeric regions. Over-expression of *Dot1* in budding yeast leads to spreading of MeK79 into telomeric regions and consequent loss of silencing by preventing the association of telomeric silencer proteins Sir2 and Sir3, which cannot bind Me79K histone H3 (van Leeuwen *et al.*, 2002).

Chromatin remodeling

Transcription in eukaryotic cells occurs in the context of chromatin, which can act as a barrier

by limiting access to *cis*-acting regulatory elements by nuclear factors. In vitro-assembled chromatin templates are refractory to transcription by general transcription factors and RNA polymerase II (Becker and Horz, 2002). Chromatin structure is perturbed in the neighborhood of expressed genes and is most obvious at promoters and enhancers, where hypersensitivity to nucleases marks sites that no longer carry canonical nucleosomes, and to which transcription factors bind. Granting access to the *cis*-acting elements in DNA, a prerequisite for any further action of the *trans*-acting factors involved, requires the establishment of local heterogeneity of chromatin and, in some cases, extensive remodeling of nucleosomal structures (Fleming and Pennings, 2001). ATP-dependent chromatin-remodeling enzymes have been identified, of which the archetype is the yeast SWI/SNF activator complex, which promotes activator binding to nucleosomal DNA at a subset of gene promoters (Becker and Horz, 2002). In mammals, diverse remodeling complexes regulate gene expression in developmentally distinct and tissue-specific patterns. Their numerous subunits (up to 11 polypeptides) can be shared between different remodeling activities and are targeted to specific loci by distinct DNA binding factors. In yeast the 2Md SWI/SNF complex stimulates the binding of transcription factors such as GAL4 to nucleosomal DNA in a reaction that requires ATP hydrolysis and causes an alteration of the path of DNA around the nucleosome, possibly due to twisting of the DNA, as well as mobilizing the nucleosomes along the DNA (Cote *et al.*, 1994). SWI/SNF homologues in mammals are involved in regulating growth control, differentiation and tumour suppression. Studies on mutant versions of *SNF5*/*SMARCB1* and *BRG1*, in primary tumors and tumor-derived cell lines suggest that they inhibit neoplasia through mis-regulation of genes involved in cellular proliferation (Klochendler-Yeivin *et al.*, 2002).

The mechanisms by which chromatin is remodeled may involve nucleosome movements, their transient unfolding or their partial or even complete disassembly. Histone modifications such as acetylation, methylation and phosphorylation are likely to be an integral part of a process that involves the ordered recruitment of chromatin modifying and remodeling complexes. Increased histone acetylation has been correlated with increased transcription, and regions of heterochromatin are generally hypoacetylated (Dillon and Festenstein, 2002). A direct link between histone acetylation and transcription activation was discovered when the yeast GCN5 transcription activator was shown to be a histone acetyltransferase (HAT) (Brownell *et al.*, 1996; Kuo *et al.*, 1996). GCN5 forms part of the ADA complex, a coactivator thought to function as an adaptor between DNA-bound activators and the basal transcriptional machinery (Marcus *et al.*, 1994).

Just as there are remodeling activities associated with the formation of active chromatin domains, there are also specific remodeling activities associated with the maintenance and formation of repressive and ''closed'' chromatin structures. For example, TUP1/SSN6 in yeast and NURD in *Drosophila* and mammals is associated with HDAC activities and is required for the formation of inactive chromatin at specific loci (Edmondson *et al.*, 1998; Watson *et al.*, 2000). The interplay between activating and repressing chromatin remodeling activities may determine overall gene expression patterns during development. For example aiming HDACs at normally active loci can lead to their repression or alternatively the presence of HATs at inactive loci can lead to their activation.

DNA modification

Only one major form of DNA modification has been identified in vertebrates: a methyl group at the 5 position of cytosine (m^5C), mainly in the context of the dinucleotide CpG (Bestor, 2000). Over 70% of CpGs are methylated in vertebrate DNA, but distinct patterns can be observed between different somatic and germline tissues (Meehan, 2003). At an early stage these inherited

patterns of methylation were hypothesized to have functional consequences and initial studies indicated that there was a correlation between the methylation status of a gene and its expression state (Groudine and Weintraub, 1981; Weintraub, 1985). For example, silent genes on the inactive X-chromosomes (Xi) are methylated at their promoter regions (Lock *et al.*, 1986; 1987). A number of DNA methyltransferase enzymes (Dnmt1, Dnmt2, Dnmt3a and Dnmt3b) have been identified by biochemical and sequence analysis. All of these contain a highly conserved C-terminal catalytic domain and, with the exception (so far) of Dnmt2, a variable N-terminal extension. This domain contains a number of protein interaction motifs that allows the different Dnmts (1, 3a and 3b) to interact with each other, HP1 and HDACs (Meehan, 2003). In addition, splice and promoter specific variants which are developmentally regulated have been identified for Dnmt1, 3a and 3b. Database screening also identified *Dnmt3L*, which lacks the conserved PC and ENV residues in motifs IV and VI of the catalytic domain. Recombinant *Dnmt3L* protein is unable to methylate DNA, but it interacts with *Dnmt3a* and *Dnmt3b* and co-localizes with these enzymes in the nuclei of transfected cells. Genetic evidence suggests that *Dnmt3L* cooperates with the Dnmt3 methyltransferases to carry out de novo methylation of maternally imprinted genes during oogenesis and early mouse embryogenesis (Bourc'his *et al.*, 2001; Hata *et al.*, 2002). Developmental patterns of methylation are highly regulated in animals, but as yet we have little understanding as to how this occurs (Stancheva and Meehan, 2000; Stancheva *et al.*, 2002). Dnmt1 co-localizes with late replicating DNA in tissue culture cells and it is thought to be required for the remethylation of hemi-methylated DNA substrates that occur after DNA replication (Easwaran *et al.*, 2004; Margot *et al.*, 2001). As such it is regarded as a maintenance methylase that propagates pre-existing patterns of methylation, however recent evidence suggests that Dnmt3a and 3b are also necessary to maintain DNA methylation patterns in somatic cells due to the low fidelity of the Dnmt1 enzyme (Liang *et al.*, 2002; Okano *et al.*, 1999).

Transcriptional repression by DNA methylation

It is widely accepted that symmetrically methylated CpGs (MeCpGs) act as a molecular tag that can recruit other proteins to DNA, namely the methyl–CpG binding domain repressor proteins (MBDs: MBD1, MBD2, MBD3, MeCP2 & Kaiso) (Hendrich and Tweedie, 2003). Other proteins in the sequence databases contain MBD type domains, such as MBD4, which is involved in DNA repair and not transcription repression (Hendrich and Tweedie, 2003). MBDs can in turn act as links between methylated DNA and multiprotein complexes (MeCP1, MeCP2/SMRT, Kaiso/Ncor) with HDAC and HMT activity that can potentially lead to the formation of a closed and condensed chromatin structure (Yoon *et al.*, 2003; Zhang *et al.*, 1999; Ruzov *et al.*, 2004). Loss of methylation, by inactivation of Dnmts, can abrogate the cascade of events mentioned above, and this situation coupled with acetylated histones leads to a more open euchromatic structure which is amenable to active gene transcription (Meehan, 2003).

Inhibition of the maintenance methyltransferase activity, Dnmt1, in amphibians and mice leads to upregulation of many genes (up to 10% of expressed sequence tags (ESTs) in mice). In amphibians there is clear evidence that DNA methylation is used to regulate developmental gene activation in early embryos. Prior to their activation at the mid-blastula transition (MBT), the promoters of the Xbra, c-Myc and TFIIA genes are methylated, whereas after the MBT, when the genes are expressed, these promoters are hypomethylated (Jaenisch and Bird, 2003; Meehan, 2003). In mammals, DNA methylation is probably not used in the same developmental context, as many promoters (up to 60%) are never methylated irrespective of their tissue specificity (e.g. the

human α-globin gene). However, DNA methylation is utilized in many silencing processes, including X chromosome inactivation, silencing of retroviruses and genomic imprinting, whereby the mono-allelic expression pattern of an autosomal gene is determined by the parental origin of the gene (Ferguson-Smith *et al.*, 2004; Hashimshony *et al.*, 2003). Recent work suggests that there are many emerging exceptions to the observation that CpG islands are not methylated during mammalian development. For example, the restriction landmark genome scanning (RLGS) technique uses the methylation sensitive restriction enzyme NotI (GCGGCCGC) to identify non-methylated CpG islands. Profiling of CpG islands in a range of rodent somatic and germ cell tissues suggests that the methylation status of up to 18% of non-X-linked CpG islands is developmentally regulated. Indeed, there are many methylated CpG islands in undifferentiated embryonic stem (ES) cells that change their profile upon differentiation (Hattori *et al.*, 2004; Shiota *et al.*, 2002). In addition, abnormal methylation patterns at gene promoters are associated with the progression of many cancers leading to the epigenetic silencing of essential cell cycle checkpoint genes that would normally participate in surveillance for cellular abnormalities (Fahrner *et al.*, 2002; Fahrner and Baylin, 2003). Dnmt1p can be purified in a complex with the pRb tumor suppressor and the pRb-associated E2F-1 transcriptional activator (Pradhan and Kim, 2002; Robertson *et al.*, 2000). Rb-mediated transcriptional repression from promoters containing E2F1-binding sites was enhanced by *Dnmt1* but without de novo methylation of the reporter construct. These types of interactions imply that Dnmts can be recruited to specific loci and participate in a dynamic pattern of gene silencing that is independent of their DNA methyltransferase activities.

In the next sections we present selected examples from a growing number of genetic diseases caused by mutations of genes required to set up and maintain chromatin regulatory states.

α-thalassemia/mental retardation syndrome; ATR-X

The X-linked disorder, α-thalassemia/mental retardation syndrome (ATR-X) is at present the best example of a chromatin-remodeling disorder (see Chapter 21). Affected individuals often also have a constellation of typical facial features: epicanthic folds, distinct nasal bridge and triangular-shaped mouths are common. Genital anomalies have been reported as well as skeletal abnormalities. The disease was mapped to the X-chromosome and subsequently to the *ATR-X* gene at Xq13.3 (Gibbons *et al.*, 1995; Gibbons and Higgs, 1996). The *ATR-X* gene encodes a predicted protein of 280 kDa featuring a PHD zinc finger motif (associated with many chromatin proteins) and an (adenosine triphosphatase) ATPase/helicase domain of the SWI/SNF type. The vast majority of clinically identified mutations in the *ATR-X* gene fall within these two motifs (PHD (~65%) and ATPase (~25%)), underscoring the importance of these two regions in ATR-X protein function. The PHD-type zinc finger domain of ATR-X most closely resembles that of the DNMT3 family of DNA methyltransferases, which have been shown to have transcriptional repressor activity (Aapola *et al.*, 2002). A further functional domain containing coiled coil motifs is proposed to lie between the PHD and ATPase/helicase domains of ATR-X, which interacts with murine heterochromatin protein 1 (mHP1) and the Polycomb protein EZH2 in yeast two-hybrid screens (Cardoso *et al.*, 1998; McDowell *et al.*, 1999).

ATR-X protein co-localizes with mHP1α to regions of condensed chromatin that probably correspond to centromeric heterochromatin. However, ATR-X is also associated with the nuclear matrix, which comprises the lamin–pore complex and an internal network of poorly characterized ribonucleoprotein filaments. Genes that are actively transcribed are enriched in nuclear matrix preparations, suggesting that the nuclear matrix may support the formation of nuclear domains for the regulation of gene expression

(Berube *et al.*, 2000). Many ATR-X mutants are associated with alterations in α-globin gene expression, which is located in chromosomal region 16pter-p13.3. However it should be noted that mild forms of ATR-X mutations are associated with mental retardation only, implying that there is a spectrum of phenotypes that may reflect ATR-X-dependent transcriptional competence during development. It is possible that the ATPase/helicase domains of ATR-X participate in chromatin remodeling complexes at specific loci, such as the α-globin gene, to ensure correct patterns of gene expression. ATR-X patients also exhibit defective methylation of the ribosomal DNA (rDNA) arrays and other sequence repetitive loci and this observation is now used in the molecular diagnosis of the disease (Gibbons *et al.*, 2000). It is not clear if these changes in rDNA methylation result in increased rRNA synthesis. In contrast, the Y-specific repeat DYZ2 is abnormally hypermethylated in ATR-X patients, suggesting a specific effect of ATR-X mutations on Y-chromosome repeats, perhaps involving recruitment of methyltransferase activity to these regions. The wide range of phenotypes in ATR-X individuals might be explained by hypomethylation of multiple specific sequences leading to the aberrant expression of a spectrum of genes during development, nevertheless global methylation levels in ATR-X patients are normal. Symptomatic comparison with the ATR-X mouse shows a very similar disease progression. ATR-X is part of a multiprotein complex, similar in size to SWI/SNF complexes, and its overexpression in mice leads to growth retardation, neural tube defects and embryonic lethality (Berube *et al.*, 2002). Generation of a line of mice lacking or with point mutations in the PHD domain may yield clues as to which genes are being misexpressed (if any) in ATR-X patients.

In mice, *Proliferation-Associated SNF2-like Gene* (*PASG* or *Lsh*), is also required to ensure that correct patterns of DNA methylation are maintained at centromeric satellite DNA sequences during embryogenesis (Dennis *et al.*, 2001; Sun *et al.*, 2004). The highly conserved PASG protein also encodes a member of the SNF2-like helicase subfamily, which can disrupt histone–DNA interactions. This suggests that remodeling activities like PASG/ATR-X may act through an effect on chromatin conformation, which regulates access of the Dnmts and HMTs to their substrates in vivo (Meehan *et al.*, 2001). Like *ATR-X*, *Lsh* is also associated with pericentromeric heterochromatin and there is an accumulation of di- and trimethylated histone 3 at lysine 4 (H3-K4me) at heterochromatin in *Lsh*-deficient cell lines. However, there is no change in the methylation status of H3-K9 or in the distribution of HP1 after *Lsh* deletion, implying independent regulatory mechanisms for H3-K4 and -K9 methylation (Yan *et al.*, 2003a; 2003b). *Lsh*-deficient embryonic fibroblast cells also exhibit a mitotic defect, with micronuclei, signs of abnormal spindle formation and centrosome amplification (Fan *et al.*, 2003). The failure to methylate DNA in the absence of *Lsh* function may allow access by H3-K4 methyltransferases to normally inactive chromatin. This may result as a consequence of delocalisation of methyl-CpG binding proteins, which has been observed in Dnmt1-deficient cell lines, suggesting that DNA and histone modification states are functionally linked and that one may reinforce the other. As we shall see, hypomethylation at centromeric regions can also occur in patients that carry a mutation in Dnmt3b. This also may lead to disturbed modification of histone proteins and contribute to chromosomal instability.

Rett syndrome – RTT

Rett syndrome is caused by a wide spectrum of mutations in the methyl-CpG binding protein MeCP2 (Amir *et al.*, 1999; Neul and Zoghbi, 2004). This multidomain chromatin-associated protein localizes to the densely methylated heterochromatin (consisting of major and minor satellite DNA) of the mouse genome (Lewis *et al.*, 1992; Meehan *et al.*, 1992). A methyl-CpG binding domain (MBD) of approximately 70 amino acids is devoted to

selective recognition of methyl CpG. A transcriptional repression domain (TRD) overlaps a nuclear localization signal and interacts with HDAC1/2 via SIN3a. This links DNA methylation with histone deacetylation and the formation of inactive chromatin. In test systems, MeCP2 can selectively silence methylated reporter templates (Nan *et al.*, 1997; 1998). Rett syndrome affects up to 1/10 000 births and the majority of patients are female, who develop normally for 6–18 months postnatally, but then suffer a developmental regression characterised by loss of speech, head growth deceleration, perturbed hand use and Tourette-like involuntary hand movements (Renieri *et al.*, 2003; Shahbazian and Zoghbi, 2002). After this initial deterioration, the condition of Rett individuals stabilizes and many survive to adulthood.

Mice lacking MeCP2 exhibit similar defects. As in humans, heterozygous female pups develop normally for about 6 months, but eventually develop breathing abnormalities, decreased movement and exhibit behavioral defects. Until recently it was believed that males mutant for MeCP2 would die prior to birth, but molecular analysis showed that the phenotype of MeCP2 mutations included males exhibiting severe encephalopathy (Lynch *et al.*, 2003). MeCP2-null male and female mice showed no initial phenotype, but both developed with a reduction in brain size and the appearance of neurological symptoms at approximately six weeks of age, including a stiff uncoordinated gait and reduced spontaneous movement between three and eight weeks of age (Chen *et al.*, 2001; Guy *et al.*, 2001; Shahbazian *et al.*, 2002). It is now known that most MeCP2 mutations occur in either the MBD or the TRD. In the former case, it is reasonable to believe that MeCP2 binding to methylated DNA would be impaired resulting in no repressor complex recruitment, and that TRD mutations would allow DNA binding but ineffective corepressor recruitment. Either of these scenarios would lead to accessible DNA and therefore active genes, although the actual in vivo targets of these genes are not known. MeCP2 is widely expressed and seems to be integral to brain function and neurogenesis, so potential targets are likely to be found in these tissue types.

However, work has shown that, unlike loss of DNA methylation, the absence of MeCP2 function is not associated with global activation of normally silent genes. Instead, only a few genes are affected in mouse and human tissues (Colantuoni *et al.*, 2001; Traynor *et al.*, 2002; Tudor *et al.*, 2002). A microarray analysis of mouse brains from MeCP2 null mice only identified subtle changes in gene expression after extensive statistical analysis. Some of the changes in gene expression were as little as 35% but were confirmed by RNase protection analysis. It could be that MeCP2 only affects a small proportion of developing cells that cannot be assessed by whole brain microarray technology. A candidate gene approach identified brain-derived neurotrophic factor (*BDNF*) as an MeCP2 target gene in mice. MeCP2 binds selectively to *BDNF* promoter III and functions to repress expression of this gene, perhaps in a developmental context (Chen *et al.*, 2003; Martinowich *et al.*, 2003). *BDNF* plays important roles in neuronal survival, development, and plasticity, and its mis-expression may underlie the molecular pathology of Rett syndrome in mammals. These findings are consistent with MeCP2 playing a role in the maintenance and modulation of neuronal maturity. In particular, MeCP2 may function as a key regulator of activity-dependent neuronal gene expression. Alternatively MeCP2 could have a transcription-independent role in the brain. There is evidence that MeCP2 has the ability to compact chromatin in vitro when present at high molar ratios (Georgel *et al.*, 2003). The Rett phenotype may be related in part to alterations in large-scale chromatin organization in target cells (Horike *et al.*, 2005). In this context, Angelman syndrome (AS) is an imprinted disorder caused by maternal deficiency of chromosome 15q11–13, methylation defects, or maternal mutation of UBE3A, encoding the ubiquitin ligase UBE3A/E6-AP (Ferguson-Smith *et al.*, 2003; Nicholls and Knepper, 2001). AS shares overlapping clinical features with RTT, including developmental delay, language impairment, seizures, and

stereotypic behaviors. It has been recently demonstrated that MeCP2 deficiency results in significant reduction of UBE3A/Ube3a and GABRB3/Gabrb3 expression in mouse cerebrum as well as in AS and RTT cerebral samples compared to controls (Samaco *et al.*, 2005). The maternally expressed gene *DLX5* also showed a loss of imprinting in lymphoblastoid cells from individuals with Rett syndrome and MeCP2-null mice (Horike *et al.*, 2005). These results implicate MeCP2 in the regulation of imprinted gene expression.

Work in *Xenopus laevis* identified an *xMeCP2* target gene by utilizing antisense morpholino oligonucleotide injection to deplete xMeCP2 activity from developing frog embryos (Stancheva *et al.*, 2003). This resulted in multiple developmental abnormalities of the head and dorsal axis indicative of an abnormal pattern of neurogenesis. The defects could be rescued by overexpression of wild type MeCP2 from human or *Xenopus*, but not by overexpression of two Rett-derived MeCP2 mutants. Examination of the expression of genes known to be essential for primary neurogenesis indicated that *xMeCP2* is involved in regulating expression of the neuronal inhibitor *xHairy2a*, a known target of the Notch/Delta signalling pathway. The *xHairy2a* promoter is bound by MeCP2 at nearby methylated CpGs and is upregulated in *xMeCP2* deficient embryos. Correct repression and activation of the *Hairy2a* gene depends on the MeCP2 interaction with SMRT corepressor complex via Sin3A. After *Notch*-mediated induction, the MeCP2/SMRT complex is released from the promoter, ensuring correct expression of *xHairy2a* and subsequent patterning of primary neurons. In the absence of MeCP2, an excess of *xHairy2a* expression is associated with a reduced number of primary neurons. This represents the first documented case of a gene that is normally repressed by MeCP2 in vivo. It is possible that rather than having a function in global regulation of methylated genes, MeCP2 has a more limited role in the precise control of a few genes analogous to transcription factors that bind a limited number of sites at gene promoters. The absence of mis-regulation at other Notch target genes, such as the Enhancer of Split-related (ESR) genes argues that repression by MeCP2 is not a uniform feature of Notch-responsive promoters. However, Notch can regulate aspects of neuronal maturation including dendritic branching, which is reduced in neurons affected by Rett syndrome, leaving open the possibility that subtle alterations in Notch signalling affecting a few target genes in neural tissue may contribute to the Rett syndrome phenotype (Redmond and Ghosh, 2001; Vetter, 2003). It is also worth noting that mutations in murine MBD containing repressor proteins (MBD2 and MBD1) also give rise to neuronal phenotypes (Hendrich *et al.*, 2001; Zhao *et al.*, 2003). This raises general questions about the precise roles of methyl-CpG binding proteins during mammalian development, which are not consistent with their putative role as global regulators of gene expression. Recently MeCP2 has been shown to have a preference for methyl-CpG that have adjacent A/T rich sequences, which are present in MeCP2-target genes and may explain selective gene regulation by MeCP2 (Klose *et al.*, 2005). Intriguingly, MeCP2 has also been implicated in regulation of RNA splicing, which may also contribute to its molecular pathology (Young *et al.*, 2005).

DNA methylation and cancer

In vertebrate somatic cells, epigenetic regulation of gene expression reinforces stable expression states at different loci. These ''expression states'' are associated with particular molecular signatures of DNA and chromatin modifications that are characteristic of active or repressed genes. The end result is that differentiated cells have a restricted transcriptome profile and a limited developmental potential. In cancer cells, this regulatory mechanism is altered such that the transcriptome profile is changed to one that promotes cancer progression and maintenance. In cell lines, selection for rapid cell division can impose further epigenetic changes.

In general, cancer cells possess aberrant patterns of hypomethylation at repeat sequences and hypermethylation at the promoters of many genes (Baylin and Herman, 2000; Bjornsson *et al.*, 2004; Feinberg *et al.*, 2002; Feinberg and Tycko, 2004; Jones and Laird, 1999). This process can give the cell a selective growth advantage akin to mutation or gene loss. Decreases in this epigenetic modification of DNA in repetitive and parasitic elements can lead to expression of these elements and a decrease in genome stability. Hypermethylation, on the other hand, is less common but seems to prefer CpG sequences usually found in simple repeat sequences but also clustered (CpG island) in the promoter regions of genes (Goel *et al.*, 2004; Kondo and Issa, 2004; Ricciardiello *et al.*, 2003). CpG island methylation is rare in normal cells and can repress transcription of the associated gene by recruiting MBDs and, in turn, repression complexes as detailed above (Ballestar *et al.*, 2003). A growing body of evidence implicates DNA methylation as an early mechanism in tumorigenesis. Promoter region hypermethylation of the retinoblastoma gene, responsible for the commonest ocular cancer in children, has in rare cases been detected in families affected by unilateral retinoblastoma (Ohtani-Fujita *et al.*, 1997). Additionally, studies of sporadic colon cancer have revealed increased methylation at the mismatch repair gene *hMLH1*. Downregulation of methylated *hMLH1* can be reversed upon treatment with the demethylating agent 5-azacytidine (Ricciardiello *et al.*, 2003). Such aberrant methylation events also fit into the Knudson two-hit model in the case of p16-I^{NK4A}, where one allele is mutated and nonfunctional with the other allele silenced by high levels of methylation (Myohanen *et al.*, 1998; Toyota *et al.*, 1999).

Epigenetic alterations in cancer are not uniform and are context-dependent. Previous studies have focused on hypermethylation and cancer, but it would be remiss to ignore decreased levels of this molecular mark, which may be of equal importance in establishing a transformed cell. A primary function of DNA methylation is to silence parasitic element transcription. In $Dnmt1^{-/-}$ ES cells, transposable elements become demethylated and re-express (Walsh *et al.*, 1998). If this process is global, which it appears to be, the mobilization of transposable elements and insertion into a coding exon, transcription interference or generation of an antisense element, could be deleterious to cellular function and stability. Another threat posed by elements ''jumping about'' the genome is homologous recombination. A number of examples of repeat sequence recombination have been shown suggesting that DNA methylation can inhibit recombination and is required for genome stability (Chen *et al.*, 1998; Eden *et al.*, 2003; Gaudet *et al.*, 2003; Guo *et al.*, 2004). Consistent with this, murine $Dnmt1^{-/-}$ ES cells exhibit a tenfold increased mutation rate involving gene rearrangements (Chen *et al.*, 1998). Mechanistically, how methylation of DNA suppresses homologous recombination events remains elusive, but a number of hypotheses have been offered to explain it including masking of the site of recombination; maintenance of a heterochromatic state; destabilization of the recombination intermediate; or direct interference with the recombination machinery.

A major breakthrough directly linking epigenetics and tumorigenesis was reported recently by Sansom *et al.* who generated an MBD2 (methyl-binding domain protein 2) knockout mouse in an existing background of tumor susceptibility to test whether a reduced ability to interpret the DNA methylation signal might alter tumorigenesis (Sansom *et al.*, 2003). $Apc^{Min/+}$ have multiple intestinal neoplasia modeled on the human cancer Familial Adenomatous Polyposis (FAP). Previous work had shown that $MBD2^{-/-}$ mice are normal despite subtle behavioral peculiarities including improper maternal nurturing (Hendrich *et al.*, 2001). It was shown that $APC^{Min/+}$ $MBD2^{-/-}$ mice outlived control $APC^{Min/+}$ mice and have ten times fewer adenomas. $APC^{Min/+}$ $MBD2^{+/-}$ survived for an intermediate amount of time indicating a dependence of tumor burden on the $MBD2^{+}$ allele.

Myeloid/lymphoid or mixed lineage leukemia

Productive chromosomal translocations involving the MLL (Myeloid/lymphoid or mixed-lineage leukaemia) gene on chromosomal region 11q23, together with up to 40 other genes have been observed in acute lymphoid leukemias and especially in acute myeloid leukemias (Ayton and Cleary, 2001). This acquired somatic mutation is associated with poor prognosis. MLL itself is a component of a chromatin complex that methylates histone H3 at lysine 4 via its SET domain (Yokoyama *et al.*, 2004). In acute leukemias, it is converted into an oncoprotein by acquisition of transcriptional effector domains following heterologous protein fusions with a variety of nuclear transcription factors, cofactors, or chromatin remodeling proteins. Despite this diversity, protein fusions appear to activate the MLL gene either by conferring constitutive transcriptional effector activity or by inducing forced MLL dimerization and oligomerization. In many human MLL leukemias, inappropriate expression of a subset of *Hox* genes, especially *HoxA9*, occurs (Armstrong *et al.*, 2002; Ferrando *et al.*, 2003; Yokoyama *et al.*, 2004). Analysis of mutant myeloid progenitor cells containing MLL fusions suggests that they have altered proliferation and survival properties. Wild type human MLL encodes a protein of almost 4000 amino acids containing multiple motifs, including AT, CxxC, PHD and SET domains. Many of the MLL translocations produce an open reading frame that includes the N-terminal 1400 residues, which excludes the *trans*-activation PHD and histone modifying SET domains but retains DNA binding (AT) and protein interaction (CxxC) domains. Genetic rearrangements of the CBP locus can underlie the molecular pathology of some acute myeloid leukemias (AML) (Sobulo *et al.*, 1997). These fusion proteins can contain histone acetyltransferase (HAT) domain(s) that may modify chromatin target genes by acetylation rather than methylation, resulting in aberrant gene regulation. Equally, acute promyelocytic leukemia translocations have been identified that generate activated fusion proteins which recruit HDAC nuclear corepressor complexes to a variety of haematopoetic lineage-specific gene promoters (Claus and Lubbert, 2003). This results in a block in differentiation, which can be alleviated by appropriate treatment with retinoic acid receptor and HDAC inhibitors.

Diseases of genomic imprinting

Mammals possess an epigenetic system thought to be important for fetus–mother nutrient transfer and normal development, termed genomic imprinting (Constancia *et al.*, 2002; Ferguson-Smith *et al.*, 2003b; Ferguson-Smith *et al.*, 2004; Reik *et al.*, 2003; Walter and Paulsen, 2003). Epigenetic guidelines are imprints laid down in germ cells (in most cases by DNA methyltransferases) governing how genes are expressed depending on their parental origin, be it maternal or paternal (Bjornsson *et al.*, 2004; Brannan and Bartolomei, 1999) (Figure 2.2). Usually the silent gamete is methylated but there are exceptions. In addition, there are DNA methylation-independent imprinting mechanisms, which are operative in extra-embryonic tissues (Lewis *et al.*, 2004; Umlauf *et al.*, 2004). The molecular mechanisms of genomic imprinting are complex, involving multiple imprinted loci, some of which exhibit developmental-specific imprints. Over 60 imprinted genes, which tend to cluster together, thus facilitating coordinate control of imprinted gene expression, have been found in mammals (Beechey, 2004). Three main steps can be visualized during mammalian imprint development; erasure, re-establishment and maintenance (molecular recognition). The first step occurs during primordial germ cell development, the second during gamete maturation and the last step in the developing embryo. The mechanism of imprint erasure is unclear but may involve active DNA demethylation and subsequent histone modifications (Lee *et al.*, 2002). Imposition of DNA methylation at imprinted regions is dependent on Dnmts,

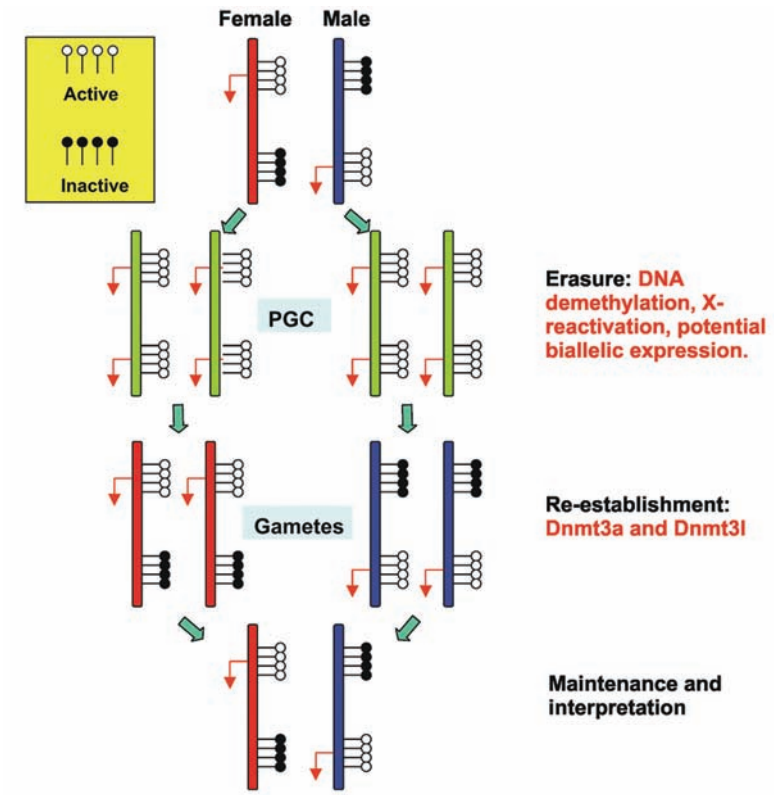

Figure 2.2 Model for genomic imprinting by DNA methylation during development. The imprints are erased during primordial germ cell (PGC) development in males and females, which allows for potential biallelic expression. During gamete formation, the imprints are re-imposed and this process is dependent on DNA methyltransferase activity (notably Dnmt3a) and the co-factor Dnmt3l. After fertilization, the parental specific imprints are maintained by DNA (Dnmt1) and histone modifying activities, resulting in allelespecific expression.

especially Dnmt3a and the cofactor Dnmt3L (Kaneda *et al.*, 2004; Hata *et al.*, 2002). The last step involves read-out of the imprinted regions and can involve selective silencing of one allele due to DNA methylation at regulatory regions (sometimes promoter regions). In many cases, mono-allelic expression is associated with antisense transcripts that are integral to silencing at multiple loci in imprinted regions (Horike *et al.*, 2000).

The association between genomic imprinting and brain disorders with influences on complex behavior is exemplified by the Prader-Willi and Angelman syndromes (PWS and AS), two neurobehavioral disorders with distinct clinical manifestations (Ferguson-Smith *et al.*, 2004; Nicholls and Knepper, 2001). Patients with PWS present with hypotonia at birth, obesity, short stature, mental retardation, hypogonadism and a characteristic facial appearance. AS is characterized by microbrachycephaly, large mouth with tongue protrusion and prognathism; mental retardation is severe with absence of speech. Both diseases are associated with deficiencies in the same region of human chromosome 15q11–q13 due to an unequal crossing over between low copy repeats. These deletions are of paternal origin in PWS

patients and of maternal origin in the case of AS. In PWS, the deletion occurs in an imprinting centre (IC) within the promoter of the imprinted gene *SNURF-SNRPN*, while in Angelman syndrome deletions are found a few thousand basepairs upstream of the IC (Imprinting Centre), which is active in germline and/or early postzygotic development (Mann and Bartolomei, 1999; Nicholls and Knepper, 2001; Yang *et al.*, 1998). The presence of the IC is integral to the establishment of allele-specific differences in DNA methylation, chromatin organization, histone modification and expression (Nicholls and Knepper, 2001). In PWS, otherwise paternally expressed genes are silenced and methylated, in contrast to AS where genes that are otherwise repressed are now demethylated and switched on. A mouse line harbouring a deletion in the putative PWS-IC region of the *SNRPN* promoter displays many human PWS-like symptoms and aberrantly silences a number of imprinted genes in this region. (Yang *et al.*, 1998). A mouse model of AS was generated by deletion of the *UBE3A* gene (Miura *et al.*, 2002). Once again, the mice displayed striking phenotypic and behavioral similarity to their human counterparts. Microarray analysis has identified candidate genes that are mis-expressed in PWS patients including *UBE3A* and *ATP10C* (Bittel *et al.*, 2003). However, PWS (unlike AS) does not appear to be due to single gene mutations. A comprehensive analysis of methylation patterns in PWS/AS families and the mouse models are needed to clarify the molecular mechanisms of these complex diseases and how these methylation patterns are interpreted by nuclear factors (see section on Rett syndrome).

Another disorder of imprinting, the overgrowth disorder Beckwith–Wiedemann syndrome (BWS) on human chromosome 11p15, has been associated with two non-adjacent chromosomal regions; a telomeric domain coincident with the *IGF2* and *H19* loci and a centromeric domain. (Ferguson-Smith *et al.*, 2004; Walter and Paulsen, 2003; Weksberg *et al.*, 2003). It has been suggested that incorrect imprinting in the telomeric region yields overgrowth and tumors, whereas those in centromeric regions involve the characteristic BWS congenital malformations (Bjornsson *et al.*, 2004; Cui *et al.*, 2002; 2003).

Two basic questions are the timing of the epigenotype spreading and the mechanism of spreading. Is spreading propagated during gametogenesis or is it post-zygotic, or both? In addition, is spreading transcription-mediated by virtue of promoters and anti-sense transcripts, or are ICs spatially arranged to come into contact with distant ICs, perhaps by matrix-attachment sites? An emerging theme is that non-coding RNAs play a role in genome modification in animals as well as plants and fungi (Aufsatz *et al.*, 2002; Volpe *et al.*, 2003). Antisense RNA transcripts are implicated in the regulation of the autosomally imprinted *Igf2r-Air* domain and in X-chromosome inactivation in female mammalian cells (Plath *et al.*, 2002; Sleutels *et al.*, 2002). The presence of multiple micro RNA genes at imprinted regions in mammals may be functionally relevant (Beechey, 2004; Seitz *et al.*, 2004). Recently it has been demonstrated that small interfering or siRNAs targeted to CpG islands of an E-cadherin promoter induced significant DNA methylation and histone H3 lysine 9 methylation in both MCF-7 and normal mammary epithelial cells, leading to gene silencing (Kawasaki and Taira, 2004).

As assisted reproductive technologies (ARTs) are increasingly used to overcome infertility, there is concern that the children conceived may be susceptible to imprinting disorders involving chromatin modification (Maher *et al.*, 2003). As yet, the evidence that ART may be associated with genetic imprinting disorders is inconclusive (Schieve *et al.*, 2004). Parenthetically, Dnmt1 overexpression causes genomic hypermethylation, loss of imprinting, and embryonic lethality (Biniszkiewicz *et al.*, 2002). In addition, somatic cell nuclear transfer often results in multiple imprinting errors (Young and Beaujean, 2004). Loss of imprinting mechanisms can also be observed in many cancers involving reactivation of the normally silent allele of a growth-promoting

gene such as *Igf2*, and/or silencing of the normally active allele of a tumour suppressor gene.

Conclusions

Recent studies have revealed that several human diseases are the result of inappropriate transcriptional de-repression. Typically, both the disease phenotype and alteration of gene expression are tissue-specific. It therefore seems likely that one or more of these diseases are caused by the deregulation of a few specific genes expressed only in particular tissues, or only under certain conditions. So far, the identification of putative target genes has remained elusive. An alternative possibility is that, at least in some cases, the targets are not protein-coding genes but instead, for example, tissue-specific centromeric heterochromatin or micro RNAs (Svoboda *et al.*, 2004; White and Allshire, 2004).

New experimental approaches, including small-molecule inhibitors, antisense oligonucleotides, and RNA interference are being developed to control gene expression. Improving the potency and specificity of these methods may expedite the development of therapeutic agents aimed at correcting the molecular defects of diseases caused by inappropriate transcriptional de-repression (Brown and Strathdee, 2002; Chung, 2002; Plass, 2002). In a recent review, Andrew Feinberg and colleagues highlighted the need for an integrated approach to disease genetics and epigenetics termed CDGE, common disease genetic epidemiology in the context of both genetic and epigenetic variation (Bjornsson *et al.*, 2004). It is based on the idea that epigenetic modifications are potential disease determinants due to their heritability in a lineage-restricted pattern during cell division and their roles in expression of sequence variation. This ''epigenomic'' approach may help explain the late onset and gradual nature of most common diseases. It may also contribute towards understanding the quantitative nature of complex traits and the role of environment in disease development, which may be missed in a sequence-based approach. For example, recent studies identify intergenerational programmed effects on both birth weight and cardiovascular disease, which may represent an epigenetic mechanism for the non-genetic inheritance of a predisposition to low birth weight and adverse cardiovascular risk across a number of generations (Drake *et al.*, 2005).

Glossary

Centromeric heterochromatin – DNA at the centromeres of chromosomes is normally in a heterochromatic state (see heterochromatin), which is probably essential for its function. In many species, the centromeric DNA is organized as large tandemly repeated satellite DNA arrays based on a short (300 bp) unique AT-rich sequence.
Chromatin – can be identified by specific dye staining of the cell nucleus. It is a complex of DNA assembled onto a protein (primarily histone) core to form the basic repeating unit of chromatin, the nucleosome. It is in the context of chromatin that genes are expressed and inherited. Chromatin occurs in two major states, euchromatin and heterochromatin, with different staining properties. Chromatin forms a network through a series of higher-order folding during cell division to eventually form the metaphase chromosomes. Chromatin complexes can be extracted from eukaryotic lysed interphase nuclei. The composition of the extracted material will depend in part on the technique used and the cell type.
Chromatin condensation/decondensation – in interphase cells, euchromatin organization is based on decondensed chromatin, which can be unravelled to varying extents down to the fundamental unit of chromatin, the nucleosome. During cell division, the interphase euchromatin fibres condense into a higher order structure called a chromosome, which forms the basis of the correct segregation of the duplicated genetic material. In daughter cells, the chromatin

decondensation and formation of the nuclear envelope appears to be necessary for nuclear activities, including transcription, RNA processing and nucleolar formation.

Cis-acting – *cis*-acting regulatory elements depend upon being physically linked to their target in DNA, RNA or protein. DNA binding factors recognize cognate recognition sequences on chromatin as an address for regulation of the linked gene.

Trans-acting factors – are diffusible gene products that can act on *cis*-acting regulatory regions in DNA, RNA and proteins.

CpG island – the mammalian genome is characterized by methylation-free sequences that contain a high density of the unmethylated dinucleotide, CpG. These CpG islands are interspersed within the CpG methylated bulk chromatin and usually contain promoter or other regulatory DNA required for active transcription of linked genes. Sixty percent of mammalian genes contain CpG islands but their frequency in the genome and their structure can differ between species. They represent open chromatin and serve as excellent landmarks in the genomic landscape.

Enhancer element – is a short region of DNA sequence that can be bound by *trans*-acting factors to enhance the transcription level of genes. An enhancer does not need to be particularly close to the genes it acts on, but it is a *cis*-acting element.

DNA methylation – in animals, DNA methylation involves the enzymatic addition of a methyl group to the 5-carbon of the cytosine pyrimidine ring by cytosine DNA methyltransferases. In humans, approximately 1% of DNA bases undergo DNA methylation in a CpG dinucleotide context. DNA methylation patterns are developmentally regulated. Abnormal methylation of CpG island containing genes in cancer cells leads to their silencing.

Dominant negative – genetic mutations can be subdivided into germline mutations, which can be passed on to progeny and somatic mutations, which are not transmitted to progeny. Mutations come in many forms and are one of the driving forces of evolution. Neutral mutations are defined as mutations whose effects do not influence the fitness of either the species or the individuals. Dominant negative (DN) mutations (also called antimorphs) have an altered gene product that acts antagonistically to the wild-type allele. These mutations usually result in an altered molecular function (often inactive) and are characterized by a dominant or semi-dominant phenotype.

Epigenetic – in biology, ''epigenetics'' is concerned with alterations in phenotype due to changes in cellular properties that are inherited, but do not represent an alteration in genotype. At the molecular level, it is associated with modification of DNA (5-methyl cytosine) and chromatin (histone modifications). Particular molecular signatures are associated with active and inactive chromatin states.

Euchromatin – euchromatin (or ''true'' chromatin) is the term used for the chromatin fraction that undergoes decondensation as the cell progresses from metaphase to interphase. Most genes are generally found in euchromatin.

Genomic imprinting – is the phenomenon whereby an autosomal gene shows selective mono-allelic expression depending on the parent of origin. Some imprinted genes are expressed from a maternally inherited chromosome and silenced on the paternal chromosome; while others show the opposite expression pattern and are only expressed from a paternally inherited chromosome. ''Imprints'' can act as a silencer or an activator for imprinted genes and are associated with differentially methylated regions in the genome.

Heterochromatin – heterochromatin was originally defined as regions of chromatin that remain condensed throughout the cell cycle. It is associated with gene inactivity and genes placed close to heterochromatin can exhibit position effect variegation (PEV).

Microarray analysis – is a high-throughput method for screening a collection of microscopic DNA spots attached to a solid surface. It is used to measure the expression levels of large numbers of

genes in different tissue or cell line samples simultaneously. Measuring gene expression using microarrays is relevant to the study of disease states and is slowly becoming a diagnostic tool.

Nuclear matrix – the exact function and veracity of the nuclear matrix is under dispute as it corresponds to the morphological structures remaining after extraction of nuclei with high salt to remove most histones and other proteins. It appears to represent a proteinaceous network of fibres as observed by electron microscopy although the composition and morphology of these structures has engendered considerable controversy and debate.

Nucleosome – the basic repeat element of chromatin is the nucleosome. The nucleosome consists of a central protein complex, the histone octamer, and 1.7 turns of DNA, about 146 base pairs, which are wrapped around the histone octamer complex. There are four different types of core histone proteins which form the octamer containing two copies each of H2A, H2B, H3 and H4. Further, there is a linker histone, H1, which contacts the exit/entry of the DNA strand on the nucleosome. The nucleosome together with histone H1 is called a chromatosome.

Parasitic elements – are mobilizable DNA sequences in the genome.

Position effect variegation (PEV) – is the most universal epigenetic gene regulation mechanism in eukaryotes (from yeast to mammals). It is the variable and reversible silencing of genes based on their chromosomal position. PEV is stable through successive cell divisions, giving rise to clones of active and inactive cells in a population.

Promoter elements – are usually located about 75–80 base pairs upstream of the initiation site for transcription. They are *cis*-acting sequences that regulate gene expression. Many genes contain TATA box motifs that are bound by the TBP *trans*-acting factor as part of the process of transcription initiation.

Real time PCR – is also called quantitative (real time) PCR, and is a method of simultaneous DNA quantification and amplification. DNA is specifically amplified by the polymerase chain reaction (PCR) in the presence of quantifiable intercalatable dyes or specially adapted primers that fluoresce upon amplification. Real-time PCR is used to quantify messenger RNA or DNA samples during development in cancer tissues and screens to identify novel interactions.

Replication timing – DNA replication or DNA synthesis is the process of copying a double-stranded DNA molecule in a cell, prior to cell division. In eukaryotes, this occurs during the S-phase of the cell cycle, preceding mitosis and meiosis. The process of replication is highly regulated and consists of three steps; initiation at replication origins, replication and termination. Replication of the human genome (6 billion nucleotide pairs) is not random. Genes are generally replicated in early S-phase followed by the rest of the genome, with heterochromatic regions replicating last. Therefore, replication can begin at some origins earlier than at others. As replication nears completion, ''bubbles'' of newly replicated DNA meet and fuse, forming two new molecules.

Restriction enzyme – a type II restriction enzyme (endonuclease) cuts double-stranded DNA at specific sequences. Digestion is dependent on the presence of co-factors and in some cases the methylation state of the DNA.

RNA polymerase II – RNA polymerase II (Pol II) transcribes DNA into RNA precursors (mRNA and most snRNA), which are subsequently processed into mature transcripts. It is a large 550 kDa complex of 12 subunits, which is regulated by a complex network of associated factors including modifying enzymes. The process of transcription can be divided into three main stages: (1) Initiation; requiring assembly of the RNA polymerase complex on the gene's promoter; (2) Elongation; the synthesis of the RNA transcript, and (3) Termination; the end of RNA synthesis and disassembly of the RNA polymerase complex.

RNAi – RNA interference was defined by Andrew Fire and colleagues who observed that small double-stranded RNAs, referred to as small

interference RNA or siRNA, can induce efficient sequence-specific silencing of gene expression by targeting the associated mRNA for degradation via the DicerRISC pathway.

siRNA – small interference RNAs (siRNA) are 21∼23-nt double-stranded RNA molecules that guide the cleavage and degradation of its target RNA leading to post-transcriptional gene silencing.

Transposon – is a mobilizable DNA sequence that can transpose to different positions within the genome of a single nucleus. Many transposons (retroposons) can generate DNA copies via an RNA intermediate molecule and these copies can integrate into new genomic sites, which may lead to the formation of new mutations. Silencing pathways can be targeted to transposons, probably for long-term genome and cellular stability.

Transcriptional repression – gene silencing can result either from indirect repression, by DNA binding factors that compete with activators for a common binding site, or from sequestration of an activator by a repressor molecule. Indirect repression by DNA methylation can be mediated through methyl CpG repressor proteins that bind methyl-CpG motifs in DNA/chromatin. Repression can also occur by direct inhibition of the basal transcription machinery through enzymatic modification of the component proteins of the initiation complex or promoter chromatin (histones).

Yeast two-hybrid screen – two-hybrid screening is a technique used to discover novel protein–protein interactions with ''bait'' protein of interest, which is tethered to the DNA binding domain of a *trans*-activating protein (Gal4). Without the activating domain (AD), the hybrid cannot activate its target (selection) gene. The AD can be provided by a ''prey'' protein that is encoded within a cDNA library which is constructed in the same reading frame as the AD region. Yeast strains containing the bait and prey are mated and only strains which contain fusion proteins that can interact with the bait will grow. In this way, potential interacting proteins are identified.

Acknowledgments

We thank Wendy Bickmore for very helpful comments and suggestions during the preparation of this review.

REFERENCES

Aapola, U., Liiv, I. and Peterson, P. (2002). Imprinting regulator DNMT3L is a transcriptional repressor associated with histone deacetylase activity. *Nucleic Acids Res.*, **30**, 3602–8.

Amir, R.E., Van, d.V., I., Wan, M. *et al.* (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat. Genet.*, **23**, 185–8.

Armstrong, S.A., Staunton, J.E., Silverman, L.B. *et al.* (2002). MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nat. Genet.*, **30**, 41–7.

Aufsatz, W., Mette, M.F., van der, W.J., Matzke, A.J. and Matzke, M. (2002). RNA-directed DNA methylation in Arabidopsis. *Proc. Natl. Acad. Sci. USA*, **99** Suppl 4, 16499–506.

Ayton, P.M. and Cleary, M.L. (2001). Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins. *Oncogene*, **20**, 5695–707.

Ballestar, E., Paz, M.F., Valle, L. *et al.* (2003). Methyl-CpG binding proteins identify novel sites of epigenetic inactivation in human cancer. *EMBO J.*, **22**, 6335–45.

Baylin, S.B. and Herman, J.G. (2000). DNA hypermethylation in tumorigenesis: epigenetics joins genetics. *Trends Genet.*, **16**, 168–74.

Becker, P.B. and Horz, W. (2002). ATP-dependent nucleosome remodeling. *Annu. Rev. Biochem.*, **71**, 247–73.

Beechey, C. (2004). http://www.mgu.har.mrc.ac.uk/research/imprinting/largemap.html.

Bentley, D.R. (2004). Genomes for medicine. *Nature*, **429**, 440–5.

Berube, N.G., Jagla, M., Smeenk, C. *et al.* (2002). Neurodevelopmental defects resulting from ATRX overexpression in transgenic mice. *Hum. Mol. Genet.*, **11**, 253–61.

Berube, N.G., Smeenk, C.A. and Picketts, D.J. (2000). Cell cycle-dependent phosphorylation of the ATRX protein correlates with changes in nuclear matrix and chromatin association. *Hum. Mol. Genet.*, **9**, 539–47.

Bestor, T. H. (2000). The DNA methyltransferases of mammals. *Hum. Mol. Genet.*, **9**, 2395–402.

Biniszkiewicz, D., Gribnau, J., Ramsahoye, B. *et al.* (2002). Dnmt1 overexpression causes genomic hypermethylation, loss of imprinting, and embryonic lethality. *Mol. Cell Biol.*, **22**, 2124–35.

Bittel, D. C., Kibiryeva, N., Talebizadeh, Z. and Butler, M. G. (2003). Microarray analysis of gene/transcript expression in Prader-Willi syndrome: deletion versus UPD. *J. Med. Genet.*, **40**, 568–74.

Bjornsson, H. T., Fallin, M. D. and Feinberg, A. P. (2004). An integrated epigenetic and genetic approach to common human disease. *Trends Genet.*, **20**, 350–8.

Bourc'his, D., Xu, G. L., Lin, C. S., Bollman, B. and Bestor, T. H. (2001). Dnmt3L and the establishment of maternal genomic imprints. *Science*, **294**, 2536–9.

Brannan, C. I. and Bartolomei, M. S. (1999). Mechanisms of genomic imprinting. *Curr. Opin. Genet. Dev.*, **9**, 164–70.

Brown, R. and Strathdee, G. (2002). Epigenomics and epigenetic therapy of cancer. *Trends Mol. Med.*, **8**, S43–S48.

Brownell, J. E., Zhou, J., Ranalli, T. *et al.* (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell*, **84**, 843–51.

Cardoso, C., Timsit, S., Villard, L. *et al.* (1998). Specific interaction between the XNP/ATR-X gene product and the SET domain of the human EZH2 protein. *Hum. Mol. Genet.*, **7**, 679–84.

Chambeyron, S. and Bickmore, W. A. (2004). Does looping and clustering in the nucleus regulate gene expression? *Curr. Opin. Cell Biol.*, **16**, 256–62.

Chen, R. Z., Akbarian, S., Tudor, M. and Jaenisch, R. (2001). Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. *Nat. Genet.*, **27**, 327–31.

Chen, R. Z., Pettersson, U., Beard, C., Jackson-Grusby, L. and Jaenisch, R. (1998). DNA hypomethylation leads to elevated mutation rates. *Nature*, **395**, 89–93.

Chen, W. G., Chang, Q., Lin, Y. *et al.* (2003). Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2. *Science*, **302**, 885–9.

Chung, D. (2002). Histone modification: the 'next wave' in cancer therapeutics. *Trends Mol. Med.*, **8**, S10–S11.

Claus, R. and Lubbert, M. (2003). Epigenetic targets in hematopoietic malignancies. *Oncogene*, **22**, 6489–96.

Colantuoni, C., Jeon, O. H., Hyder, K. *et al.* (2001). Gene expression profiling in postmortem Rett Syndrome brain: differential gene expression and patient classification. *Neurobiol. Dis.*, **8**, 847–65.

Constancia, M., Hemberger, M., Hughes, J. *et al.* (2002). Placental-specific IGF-II is a major modulator of placental and fetal growth. *Nature*, **417**, 945–8.

Cote, J., Quinn, J., Workman, J. L. and Peterson, C. L. (1994). Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. *Science*, **265**, 53–60.

Cowell, I. G., Aucott, R., Mahadevaiah, S. K. *et al.* (2002). Heterochromatin, HP1 and methylation at lysine 9 of histone H3 in animals. *Chromosoma*, **111**, 22–36.

Cui, H., Cruz-Correa, M., Giardiello, F. M. *et al.* (2003). Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. *Science*, **299**, 1753–5.

Cui, H., Onyango, P., Brandenburg, S. *et al.* (2002). Loss of imprinting in colorectal cancer linked to hypomethylation of H19 and IGF2. *Cancer Res.*, **62**, 6442–6.

Davey, C., Pennings, S., Meersseman, G., Wess, T. J. and Allan, J. (1995). Periodicity of strong nucleosome positioning sites around the chicken adult beta-globin gene may encode regularly spaced chromatin. *Proc. Natl. Acad. Sci. USA*, **92**, 11210–14.

Dennis, K., Fan, T., Geiman, T., Yan, Q. and Muegge, K. (2001). Lsh, a member of the SNF2 family, is required for genome-wide methylation. *Genes Dev.*, **15**, 2940–4.

Dillon, N. and Festenstein, R. (2002). Unravelling heterochromatin: competition between positive and negative factors regulates accessibility. *Trends Genet.*, **18**, 252–8.

Drake, A. J., Walker, B. R. and Seckl, J. R. (2005). Intergenerational consequences of fetal programming by in utero exposure to glucocorticoids in rats. *Am. J. Physiol Regul. Integr. Comp Physiol.*, **288**, R34–R38.

Easwaran, H. P., Schermelleh, L., Leonhardt, H. and Cardoso, M. C. (2004). Replication-independent chromatin loading of Dnmt1 during G2 and M phases. *EMBO Rep.*, **5**, 1181–6.

Eden, A., Gaudet, F., Waghmare, A. and Jaenisch, R. (2003). Chromosomal instability and tumors promoted by DNA hypomethylation. *Science*, **300**, 455.

Edmondson, D. G., Zhang, W., Watson, A. *et al.* (1998). In vivo functions of histone acetylation/deacetylation in Tup1p repression and Gcn5p activation. *Cold Spring Harb. Symp. Quant. Biol.*, **63**, 459–68.

Fahrner, J. A. and Baylin, S. B. (2003). Heterochromatin: stable and unstable invasions at home and abroad. *Genes Dev.*, **17**, 1805–12.

Fahrner, J. A., Eguchi, S., Herman, J. G. and Baylin, S. B. (2002). Dependence of histone modifications and gene

expression on DNA hypermethylation in cancer. *Cancer Res.*, **62**, 7213–18.

Fan, T., Yan, Q., Huang, J. *et al.* (2003). Lsh-deficient murine embryonal fibroblasts show reduced proliferation with signs of abnormal mitosis. *Cancer Res.*, **63**, 4677–83.

Feinberg, A. P., Oshimura, M. and Barrett, J. C. (2002). Epigenetic mechanisms in human disease. *Cancer Res.*, **62**, 6784–7.

Feinberg, A. P. and Tycko, B. (2004). The history of cancer epigenetics. *Nat. Rev. Cancer*, **4**, 143–53.

Ferguson-Smith, A., Lin, S. P., Tsai, C. E., Youngson, N. and Tevendale, M. (2003). Genomic imprinting—insights from studies in mice. *Semin. Cell Dev. Biol.*, **14**, 43–9.

Ferguson-Smith, A. C., Lin, S. P. and Youngson, N. (2004). Regulation of gene activity and repression: a consideration of unifying themes. *Curr. Top. Dev. Biol.*, **60**, 197–213.

Ferrando, A. A., Armstrong, S. A., Neuberg, D. S. *et al.* (2003). Gene expression signatures in MLL-rearranged T-lineage and B-precursor acute leukemias: dominance of HOX dysregulation. *Blood*, **102**, 262–8.

Fire, A., Xu, S., Montgomery, M. K. *et al.* (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, **391**, 806–11.

Fleming, A. B. and Pennings, S. (2001). Antagonistic remodelling by Swi-Snf and Tup1-Ssn6 of an extensive chromatin region forms the background for FLO1 gene regulation. *EMBO J.*, **20**, 5219–31.

Gaudet, F., Hodgson, J. G., Eden, A. *et al.* (2003). Induction of tumors in mice by genomic hypomethylation. *Science*, **300**, 489–92.

Georgel, P. T., Horowitz-Scherer, R. A., Adkins, N. *et al.* (2003). Chromatin compaction by human MeCP2. Assembly of novel secondary chromatin structures in the absence of DNA methylation. *J. Biol. Chem.*, **278**, 32181–8.

Gibbons, R. J. and Higgs, D. R. (1996). The alpha-thalassemia/mental retardation syndromes. *Medicine (Baltimore)*, **75**, 45–52.

Gibbons, R. J., McDowell, T. L., Raman, S. *et al.* (2000). Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation. *Nat. Genet.*, **24**, 368–71.

Gibbons, R. J., Picketts, D. J., Villard, L. and Higgs, D. R. (1995). Mutations in a putative global transcriptional regulator cause X-linked mental retardation with alpha-thalassemia (ATR-X syndrome). *Cell*, **80**, 837–45.

Gilbert, N., Boyle, S., Fiegler, H. *et al.* (2004). Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. *Cell*, **118**, 555–66.

Gilbert, N., Boyle, S., Sutherland, H. *et al.* (2003). Formation of facultative heterochromatin in the absence of HP1. *EMBO J.*, **22**, 5540–50.

Goel, A., Arnold, C. N., Niedzwiecki, D. *et al.* (2004). Frequent inactivation of PTEN by promoter hypermethylation in microsatellite instability-high sporadic colorectal cancers. *Cancer Res.*, **64**, 3014–21.

Groudine, M. and Weintraub, H. (1981). Activation of globin genes during chicken development. *Cell*, **24**, 393–401.

Guo, G., Wang, W. and Bradley, A. (2004). Mismatch repair genes identified using genetic screens in Blm-deficient embryonic stem cells. *Nature*, **429**, 891–5.

Guy, J., Hendrich, B., Holmes, M., Martin, J. E. and Bird, A. (2001). A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. *Nat. Genet.*, **27**, 322–6.

Hashimshony, T., Zhang, J., Keshet, I., Bustin, M. and Cedar, H. (2003). The role of DNA methylation in setting up chromatin structure during development. *Nat. Genet.*, **34**, 187–92.

Hata, K., Okano, M., Lei, H. and Li, E. (2002). Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development*, **129**, 1983–93.

Hattori, N., Abe, T., Hattori, N. *et al.* (2004). Preference of DNA methyltransferases for CpG islands in mouse embryonic stem cells. *Genome Res.*, **14**, 1733–40.

Hendrich, B., Guy, J., Ramsahoye, B., Wilson, V. A. and Bird, A. (2001). Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. *Genes Dev.*, **15**, 710–23.

Hendrich, B. and Tweedie, S. (2003). The methyl-CpG binding domain and the evolving role of DNA methylation in animals. *Trends Genet.*, **19**, 269–77.

Horike, S., Cai, S., Miyano, M., Cheng, J. F. and Kohwi-Shigematsu, T. (2005). Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. *Nat. Genet.*, **37**, 31–40.

Horike, S., Mitsuya, K., Meguro, M. *et al.* (2000). Targeted disruption of the human LIT1 locus defines a putative imprinting control element playing an essential role in Beckwith–Wiedemann syndrome. *Hum. Mol. Genet.*, **9**, 2075–83.

Huang, C., Sloan, E. A. and Boerkoel, C. F. (2003). Chromatin remodeling and human disease. *Curr. Opin. Genet. Dev.*, **13**, 246–52.

Jaenisch, R. and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.*, **33** Suppl, 245–54.

Jenuwein, T. and Allis, C. D. (2001). Translating the histone code. *Science*, **293**, 1074–80.

Johnston, M. V., Alemi, L. and Harum, K. H. (2003). Learning, memory, and transcription factors. *Pediatr. Res.*, **53**, 369–74.

Jones, P. A. and Laird, P. W. (1999). Cancer epigenetics comes of age. *Nat. Genet.*, **21**, 163–7.

Kaneda, M., Okano, M., Hata, K. *et al.* (2004). Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature*, **429**, 900–3.

Kawasaki, H. and Taira, K. (2004). Induction of DNA methylation and gene silencing by short interfering RNAs in human cells. *Nature*, **431**, 211–17.

Khorasanizadeh, S. (2004). The nucleosome: from genomic organization to genomic regulation. *Cell*, **116**, 259–72.

Klochendler-Yeivin, A., Muchardt, C. and Yaniv, M. (2002). SWI/SNF chromatin remodeling and cancer. *Curr. Opin. Genet. Dev.*, **12**, 73–9.

Klose, R. J., Sarraf, S. A., Schmiedeberg, L. *et al.* (2005). DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG. *Mol. Cell*, **19**, 667–78.

Kondo, Y. and Issa, J. P. (2004). Epigenetic changes in colorectal cancer. *Cancer Metastasis Rev.*, **23**, 29–39.

Kornberg, R. D. and Lorch, Y. (1999). Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell*, **98**, 285–94.

Kuo, M. H., Brownell, J. E., Sobel, R. E. *et al.* (1996). Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature*, **383**, 269–72.

Lachner, M., O'Sullivan, R. J. and Jenuwein, T. (2003). An epigenetic road map for histone lysine methylation. *J. Cell Sci.*, **116**, 2117–24.

Lee, J., Inoue, K., Ono, R. *et al.* (2002). Erasing genomic imprinting memory in mouse clone embryos produced from day 11.5 primordial germ cells. *Development*, **129**, 1807–17.

Lehnertz, B., Ueda, Y., Derijck, A. A. *et al.* (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol.*, **13**, 1192–200.

Lewis, A., Mitsuya, K., Umlauf, D. *et al.* (2004). Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of DNA methylation. *Nat. Genet.*, **36**, 1291–5.

Lewis, J. D., Meehan, R. R., Henzel, W. J. *et al.* (1992). Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell*, **69**, 905–14.

Liang, G., Chan, M. F., Tomigahara, Y. *et al.* (2002). Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. *Mol. Cell Biol.*, **22**, 480–91.

Lock, L. F., Melton, D. W., Caskey, C. T. and Martin, G. R. (1986). Methylation of the mouse hprt gene differs on the active and inactive X chromosomes. *Mol. Cell Biol.*, **6**, 914–24.

Lock, L. F., Takagi, N. and Martin, G. R. (1987). Methylation of the Hprt gene on the inactive X occurs after chromosome inactivation. *Cell*, **48**, 39–46.

Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. and Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature*, **389**, 251–60.

Luger, K. and Richmond, T. J. (1998). The histone tails of the nucleosome. *Curr. Opin. Genet. Dev.*, **8**, 140–6.

Lynch, S. A., Whatley, S. D., Ramesh, V., Sinha, S. and Ravine, D. (2003). Sporadic case of fatal encephalopathy with neonatal onset associated with a T158M missense mutation in MECP2. *Arch. Dis. Child Fetal Neonatal Ed*, **88**, F250–F252.

Maher, E. R., Afnan, M. and Barratt, C. L. (2003). Epigenetic risks related to assisted reproductive technologies: epigenetics, imprinting, ART and icebergs? *Hum. Reprod.*, **18**, 2508–11.

Maher, E. R., Brueton, L. A., Bowdin, S. C. *et al.* (2003*b*). Beckwith-Wiedemann syndrome and assisted reproduction technology (ART). *J. Med. Genet.*, **40**, 62–4.

Maison, C. and Almouzni, G. (2004). HP1 and the dynamics of heterochromatin maintenance. *Nat. Rev. Mol. Cell Biol.*, **5**, 296–304.

Mann, M. R. and Bartolomei, M. S. (1999). Towards a molecular understanding of Prader-Willi and Angelman syndromes. *Hum. Mol. Genet.*, **8**, 1867–73.

Marcus, G. A., Silverman, N., Berger, S. L., Horiuchi, J. and Guarente, L. (1994). Functional similarity and physical association between GCN5 and ADA2: putative transcriptional adaptors. *EMBO J.*, **13**, 4807–15.

Margot, J. B., Cardoso, M. C. and Leonhardt, H. (2001). Mammalian DNA methyltransferases show different subnuclear distributions. *J. Cell Biochem.*, **83**, 373–9.

Marmorstein, R. (2003). Structure of SET domain proteins: a new twist on histone methylation. *Trends Biochem. Sci.*, **28**, 59–62.

Martinowich, K., Hattori, D., Wu, H. *et al.* (2003). DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. *Science*, **302**, 890–3.

McDowell, T. L., Gibbons, R. J., Sutherland, H. *et al.* (1999). Localization of a putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes. *Proc. Natl. Acad. Sci. USA*, **96**, 13983–8.

Meehan, R. R. (2003). DNA methylation in animal development. *Semin. Cell Dev. Biol.*, **14**, 53–65.

Meehan, R. R., Kao, C. F. and Pennings, S. (2003). HP1 binding to native chromatin in vitro is determined by the hinge region and not by the chromodomain. *EMBO J.*, **22**, 3164–74.

Meehan, R. R., Lewis, J. D. and Bird, A. P. (1992). Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA. *Nucleic Acids Res.*, **20**, 5085–92.

Meehan, R. R., Pennings, S. and Stancheva, I. (2001). Lashings of DNA methylation, forkfuls of chromatin remodeling. *Genes Dev.*, **15**, 3231–6.

Meersseman, G., Pennings, S. and Bradbury, E. M. (1992). Mobile nucleosomes—a general behavior. *EMBO J.*, **11**, 2951–9.

Mermoud, J. E., Popova, B., Peters, A. H., Jenuwein, T. and Brockdorff, N. (2002). Histone H3 lysine 9 methylation occurs rapidly at the onset of random X chromosome inactivation. *Curr. Biol.*, **12**, 247–51.

Miura, K., Kishino, T., Li, E. *et al.* (2002). Neurobehavioral and electroencephalographic abnormalities in Ube3a maternal-deficient mice. *Neurobiol. Dis.*, **9**, 149–59.

Myohanen, S. K., Baylin, S. B. and Herman, J. G. (1998). Hypermethylation can selectively silence individual p16ink4A alleles in neoplasia. *Cancer Res.*, **58**, 591–3.

Nan, X., Campoy, F. J. and Bird, A. (1997). MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell*, **88**, 471–81.

Nan, X., Ng, H. H., Johnson, C. A. *et al.* (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*, **393**, 386–9.

Neul, J. L. and Zoghbi, H. Y. (2004). Rett syndrome: a prototypical neurodevelopmental disorder. *Neuroscientist*, **10**, 118–28.

Nicholls, R. D. and Knepper, J. L. (2001). Genome organization, function, and imprinting in Prader-Willi and Angelman syndromes. *Annu. Rev. Genomics Hum. Genet.*, **2**, 153–75.

Nielsen, S. J., Schneider, R., Bauer, U. M. *et al.* (2001). Rb targets histone H3 methylation and HP1 to promoters. *Nature*, **412**, 561–5.

Ohtani-Fujita, N., Dryja, T. P., Rapaport, J. M. *et al.* (1997). Hypermethylation in the retinoblastoma gene is associated with unilateral, sporadic retinoblastoma. *Cancer Genet. Cytogenet.*, **98**, 43–9.

Okano, M., Bell, D. W., Haber, D. A. and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*, **99**, 247–57.

Pennings, S., Allan, J. and Davey, C. (2005). DNA methylation, nucleosome formation and positioning. *Brief. Funct. Genomic. Proteomic* , **3**, 351–61.

Pennisi, E. (2003). Human genome. A low number wins the GeneSweep Pool. *Science*, **300**, 1484.

Plass, C. (2002). Cancer epigenomics. *Hum. Mol. Genet.*, **11**, 2479–88.

Plath, K., Mlynarczyk-Evans, S., Nusinow, D. A. and Panning, B. (2002). Xist RNA and the mechanism of X chromosome inactivation. *Annu. Rev. Genet.*, **36**, 233–78.

Pradhan, S. and Kim, G. D. (2002). The retinoblastoma gene product interacts with maintenance human DNA (cytosine-5) methyltransferase and modulates its activity. *EMBO J.*, **21**, 779–88.

Quivy, J. P., Roche, D., Kirschner, D. *et al.* (2004). A CAF-1 dependent pool of HP1 during heterochromatin duplication. *EMBO J*, **23**, 3516–26.

Rea, S., Eisenhaber, F., O'Carroll, D. *et al.* (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature*, **406**, 593–9.

Redmond, L. and Ghosh, A. (2001). The role of Notch and Rho GTPase signaling in the control of dendritic development. *Curr. Opin. Neurobiol.*, **11**, 111–17.

Reik, W., Constancia, M., Fowden, A. *et al.* (2003). Regulation of supply and demand for maternal nutrients in mammals by imprinted genes. *J. Physiol.*, **547**, 35–44.

Renieri, A., Meloni, I., Longo, I. *et al.* (2003). Rett syndrome: the complex nature of a monogenic disease. *J. Mol. Med.*, **81**, 346–54.

Ricciardiello, L., Goel, A., Mantovani, V. *et al.* (2003). Frequent loss of hMLH1 by promoter hypermethylation leads to microsatellite instability in adenomatous

polyps of patients with a single first-degree member affected by colon cancer. *Cancer Res.*, **63**, 787–92.

Richards, E. J. and Elgin, S. C. (2002). Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. *Cell*, **108**, 489–500.

Robertson, K. D., Ait-Si-Ali, S., Yokochi, T. (2000). DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. *Nat. Genet.*, **25**, 338–342.

Rodriguez-Melendez, R. and Zempleni, J. (2003). Regulation of gene expression by biotin (review). *J. Nutr. Biochem.*, **14**, 680–90.

Ruzov, A., Dunican, D. S., Prokhortchouk, A. *et al.* (2004). Kaiso is a genome-wide repressor of transcription that is essential for amphibian development. *Development*, **131**, 6185–94.

Samaco, R. C., Hogart, A. and Lasalle, J. M. (2005). Epigenetic overlap in autism-spectrum neurodevelopmental disorders: MECP2 deficiency causes reduced expression of UBE3A and GABRB3. *Hum. Mol. Genet*, **14**, 483–92.

Sansom, O. J., Berger, J., Bishop, S. M. *et al.* (2003). Deficiency of Mbd2 suppresses intestinal tumorigenesis. *Nat. Genet.*, **34**, 145–7.

Schieve, L. A., Rasmussen, S. A., Buck, G. M. *et al.* (2004). Are children born after assisted reproductive technology at increased risk for adverse health outcomes? *Obstet. Gynecol.*, **103**, 1154–63.

Seitz, H., Royo, H., Bortolin, M. L. *et al.* (2004). A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. *Genome Res*, **14**, 1741–8.

Shahbazian, M., Young, J., Yuva-Paylor, L. *et al.* (2002). Mice with truncated MeCP2 recapitulate many Rett syndrome features and display hyperacetylation of histone H3. *Neuron*, **35**, 243–54.

Shahbazian, M. D. and Zoghbi, H. Y. (2002). Rett syndrome and MeCP2: linking epigenetics and neuronal function. *Am. J. Hum. Genet.*, **71**, 1259–72.

Shiota, K., Kogo, Y., Ohgane, J. *et al.* (2002). Epigenetic marks by DNA methylation specific to stem, germ and somatic cells in mice. *Genes Cells*, **7**, 961–9.

Sleutels, F., Zwart, R. and Barlow, D. P. (2002). The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature*, **415**, 810–13.

Sobulo, O. M., Borrow, J., Tomek, R. *et al.* (1997). MLL is fused to CBP, a histone acetyltransferase, in therapy-related acute myeloid leukemia with a t(11;16) (q23;p13.3). *Proc. Natl. Acad. Sci. USA*, **94**, 8732–7.

Stancheva, I., Collins, A. L., Van, d. V., I, Zoghbi, H. and Meehan, R. R. (2003). A mutant form of MeCP2 protein associated with human Rett syndrome cannot be displaced from methylated DNA by notch in Xenopus embryos. *Mol. Cell*, **12**, 425–35.

Stancheva, I., El Maarri, O., Walter, J., Niveleau, A. and Meehan, R. R. (2002). DNA methylation at promoter regions regulates the timing of gene activation in Xenopus laevis embryos. *Dev. Biol.*, **243**, 155–65.

Stancheva, I. and Meehan, R. R. (2000). Transient depletion of xDnmt1 leads to premature gene activation in Xenopus embryos. *Genes Dev.*, **14**, 313–27.

Sun, L. Q., Lee, D. W., Zhang, Q. *et al.* (2004). Growth retardation and premature aging phenotypes in mice with disruption of the SNF2-like gene, PASG. *Genes Dev.*, **18**, 1035–46.

Svoboda, P., Stein, P., Anger, M. *et al.* (2004). RNAi and expression of retrotransposons MuERV-L and IAP in preimplantation mouse embryos. *Dev. Biol.*, **269**, 276–85.

Toyota, M., Ho, C., Ahuja, N. *et al.* (1999). Identification of differentially methylated sequences in colorectal cancer by methylated CpG island amplification. *Cancer Res.*, **59**, 2307–12.

Traynor, J., Agarwal, P., Lazzeroni, L. and Francke, U. (2002). Gene expression patterns vary in clonal cell cultures from Rett syndrome females with eight different MECP2 mutations. *BMC. Med. Genet.*, **3**, 12.

Tudor, M., Akbarian, S., Chen, R. Z. and Jaenisch, R. (2002). Transcriptional profiling of a mouse model for Rett syndrome reveals subtle transcriptional changes in the brain. *Proc. Natl. Acad. Sci. USA*, **99**, 15536–41.

Turner, B. M. (2000). Histone acetylation and an epigenetic code. *Bioessays*, **22**, 836–45.

Umlauf, D., Goto, Y., Cao, R. *et al.* (2004). Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nat. Genet.*, **36**, 1296–1300.

van Leeuwen, F., Gafken, P. R. and Gottschling, D. E. (2002). Dot1p modulates silencing in yeast by methylation of the nucleosome core. *Cell*, **109**, 745–56.

Vetter, M. L. (2003). Methylation gets SMRT. Functional insights into Rett syndrome. *Dev. Cell*, **5**, 359–60.

Volpe, T., Schramke, V., Hamilton, G. L. *et al.* (2003). RNA interference is required for normal centromere function in fission yeast. *Chromosome. Res.*, **11**, 137–46.

Waddington, C. (1957). *The strategy of the genes*. London: Allen & Unwin.

Walsh, C. P., Chaillet, J. R. and Bestor, T. H. (1998). Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat. Genet.*, **20**, 116–17.

Walter, J. and Paulsen, M. (2003). Imprinting and disease. *Semin. Cell Dev. Biol.*, **14**, 101–10.

Watson, A. D., Edmondson, D. G., Bone, J. R. *et al.* (2000). Ssn6-Tup1 interacts with class I histone deacetylases required for repression. *Genes Dev.*, **14**, 2737–44.

Weintraub, H. (1985). Assembly and propagation of repressed and depressed chromosomal states. *Cell*, **42**, 705–11.

Weksberg, R., Smith, A. C., Squire, J. and Sadowski, P. (2003). Beckwith-Wiedemann syndrome demonstrates a role for epigenetic control of normal development. *Hum. Mol. Genet.*, **12** *Spec No 1*, R61–R68.

White, S. A. and Allshire, R. C. (2004). Loss of Dicer fowls up centromeres. *Nat. Cell Biol.*, **6**, 696–7.

Wolffe, A. P. (1998). *Chromatin: structure and function*. Academic Press, London.

Wyrick, J. J., Holstege, F. C., Jennings, E. G. *et al.* (1999). Chromosomal landscape of nucleosome-dependent gene expression and silencing in yeast. *Nature*, **402**, 418–21.

Yan, Q., Cho, E., Lockett, S. and Muegge, K. (2003*a*). Association of Lsh, a regulator of DNA methylation, with pericentromeric heterochromatin is dependent on intact heterochromatin. *Mol. Cell Biol.*, **23**, 8416–28.

Yan, Q., Huang, J., Fan, T., Zhu, H. and Muegge, K. (2003*b*). Lsh, a modulator of CpG methylation, is crucial for normal histone methylation. *EMBO J.*, **22**, 5154–62.

Yang, T., Adamson, T. E., Resnick, J. L. *et al.* (1998). A mouse model for Prader-Willi syndrome imprinting-centre mutations. *Nat. Genet.*, **19**, 25–31.

Yokoyama, A., Wang, Z., Wysocka, J. *et al.* (2004). Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. *Mol. Cell Biol.*, **24**, 5639–49.

Yoon, H. G., Chan, D. W., Reynolds, A. B., Qin, J. and Wong, J. (2003). N-CoR mediates DNA methylation-dependent repression through a methyl CpG binding protein Kaiso. *Mol. Cell*, **12**, 723–34.

Young, J. I., Hong, E. P., Castle, J. C. *et al.* (2005). Regulation of RNA splicing by the methylation-dependent transcriptional repressor methyl-CpG binding protein 2. *Proc. Natl. Acad. Sci. USA*, **102**, 17551–8.

Young, L. E. and Beaujean, N. (2004). DNA methylation in the preimplantation embryo: the differing stories of the mouse and sheep. *Anim Reprod. Sci.*, **82–83**, 61–78.

Zhang, X., Yang, Z., Khan, S. I. *et al.* (2003). Structural basis for the product specificity of histone lysine methyltransferases. *Mol. Cell*, **12**, 177–85.

Zhang, Y., Ng, H. H., Erdjument-Bromage, H. *et al.* (1999). Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev.*, **13**, 1924–35.

Zhao, X., Ueba, T., Christie, B. R. *et al.* (2003). Mice lacking methyl-CpG binding protein 1 have deficits in adult neurogenesis and hippocampal function. *Proc. Natl. Acad. Sci. USA*, **100**, 6777–82.

3

Population genetics and disease

Donald F. Conrad and Jonathan K. Pritchard

Introduction

The field of population genetics may be broadly defined as the study of the generation and maintenance of genetic variation within populations. Population genetic theory plays an important role in shaping our understanding of human genetic variation in general, and the genetic basis of common disease in particular. It also plays a central role in association- and linkage disequilibrium-based approaches to disease mapping, as these can only be properly understood within a population genetic framework.

One way to think about the role of population genetics in the study of complex disease is as follows. If it were possible to sequence every base pair of every person in a study population, conventional statistical methods would, arguably, be sufficient to make inferences about which sequence variants are associated with disease. Population genetics provides an analytical framework for predicting the nature of unobserved variation that lies between genotyped sites, or in unsampled individuals. Similarly, population genetic approaches are used to explore plausible models of complex disease, as there are at present few empirical data on the genetic basis of complex diseases. Lastly, population genetics allows us to measure the effect of genetic variation on health in an indirect way, by detecting selective effects that may be too subtle to observe directly in prospective data, or that may have affected humans during our evolutionary history. In short, biological properties of a species which we cannot directly observe are illuminated by population genetics. Human genetics, as a field of genetics that lacks the benefit of an experimental system, therefore has a special reliance on population genetic theory.

The founders of population genetics applied the existing conceptual framework of the physical sciences to their emerging field, and conceived of the collection of genomes within a population as a system. The state of the system, characterized by measurable properties such as allele, haplotype, and genotype frequencies, is shaped by ''forces'', among which are mutation, drift, and selection. When the forces are held constant long enough, the system may reach an equilibrium with regard to some or all of its measured properties. It was hoped that by understanding these forces and how they interact, a set of laws would emerge that describe our genetic world in much the same way as useful laws in physics and chemistry. Indeed, the field of population genetics arguably has one of the richest and best-developed theoretical frameworks of any area within the biological sciences. In this chapter, we provide a brief introduction to some of the central concepts in this field, as they relate to the study of complex traits. Readers who wish to pursue concepts from this chapter in greater detail may wish to consult textbooks such as Hartl and Clark (1997); Falconer and Mackay (1996); Gillespie (2004), or the various reviews referenced in this chapter.

Genetic variation and genetic disease

Population measurement of the prevalence of genetic disease is challenging due to the difficulty of accurate case ascertainment. Some mutations are incompatible with pregnancy, and many create phenotypes that are under-reported or unrecognized due to mild phenotype. The British Columbia Health Surveillance Registry has monitored over 1 million live births in Canada since 1952, and has some of the most thorough records on the incidence of genetic disease (Baird *et al.*, 1988). According to their records it is estimated that, on average, at least 53 out of 1000 newborn children will manifest a disease with a genetic component by age 25 (Table 3.1).

These numbers indicate that a substantial number of individuals in this population carry deleterious mutations that may impair their (or their children's) quality of life and possibly prevent reproduction. Why are these deleterious mutations at this particular frequency? What is the likely frequency of genetic disease in other populations? Population genetic theory aims to provide insight into these questions.

The human genome is incredibly plastic. Although mutation at some points in the genome is lethal, the genome of a newly born human may carry as many as 100 de novo mutations (Kondrashov, 1995), most of which are neutral or at worst mildly deleterious. This new variation can take the forms of small insertions/deletions, single nucleotide changes (SNPs) or, often more seriously, transposable element insertion, aneuploidy (extra chromosomes or pieces of chromosome), chromosomal rearrangements, or major duplications and deletions. We are only just beginning to understand the extent to which large-scale structural changes to the genome can be tolerated and persist at the population level (Sebat *et al.*, 2004; Iafrate *et al.*, 2004; Tuzun *et al.*, 2005; Conrad *et al.*, 2006). Interspecies comparisons between human and chimpanzee, and the direct observation of germline mutation estimate an average genomic mutation rate of around $= 2 \times 10^{-8}$ per nucleotide per generation (Nachman and Crowell, 2000; Kondrashov, 2003). It is worth noting that the human mutation rate shows age- and sex-specific effects: in general, gametic mutations are more likely in older individuals, and the mutation rate in males is higher than in females (Crow, 2000; Haldane, 1947).

A major aim of population genetics is to understand how the variation created by mutation is organized among individuals and between populations. By analyzing 1.2 billion bases of sequence data from a single population, the International SNP Map Working Group estimated that the average frequency of differences between two haploid genomes is 7.51×10^{-4}, or equivalently that a typical individual is heterozygous at one SNP per 1331 bp (Sachidanandam *et al.*, 2001). This measure of variation (sometimes called ''heterozygosity'' or ''nucleotide diversity'') was smaller on the sex chromosomes, possibly due to a smaller

Table 3.1. Nature and frequency of genetic disease

Type of genetic disease	Frequency by 25 y.o.	Frequency after 25 y.o	Lifetime frequency
Chromosomal	1.8/1000	2/1000	3.8/1000[a]
Single gene	3.6/1000	16.4/1000	20/1000[b]
Multifactorial	46.4/1000	600/1000	646.4/1000
Somatic cell	–	240/1000	240/1000[c]

Source: (Rimoin *et al.*, 1997).

[a] Includes sex chromosome aneuploidy.

[b] Includes mitochondrial gene mutations.

[c] Does not include inherited single gene cancer syndromes.

Table 3.2. Frequency of Mendelian mutation classes

Mutation type	Number of entries
Single base-pair substitutions	
Missense/nonsense	29 525
Splicing	4910
Regulatory	641
Other lesions	
Micro-deletions (≤20 bp)	8477
Micro-insertions (≤20 bp)	3763
Indels (≤20 bp)	722
Gross (>20 bp) deletions	2799
Gross (>20 bp) duplications and insertions	448
Complex rearrangements (including inversions)	368
Repeat variations	132
Total	51 385

Source: (HGMD, November 2003; Stenson *et al.*, 2003).

effective population size or increased selective pressure. Genome-wide estimates of variation can be informative about demographic and selective processes acting on populations, but what amount of the variation that we observe is translated into biological differences between individuals? Of the SNPs reported by Sachidanandam *et al.* (2001), just 4.2% were identified in exons. Subsequent studies have estimated that 50.3% of exonic SNPs result in amino acid replacements, 49% are silent, and 0.7% produce a stop codon (Salisbury *et al.*, 2003).

By analyzing the relative frequency of synonymous and non-synonymous amino acid polymorphism in human populations, Fay *et al.* (2001) estimated that the average person carries 500–1200 mildly deleterious mutations. The Human Gene Mutation Database (HGMD) lists over 50 000 mutations in 2000 genes leading to disease (mostly Mendelian), and is a useful empirical tool for understanding the physical nature of deleterious mutation. About 7% of the changes listed in HGMD are insertions or duplications, 22% are deletions, and 59% are single-nucleotide missense or nonsense mutations (Stenson *et al.*, 2003). A summary of all entries in HGMD is shown in Table 3.2. Botstein and Risch (2003) have recently argued (somewhat controversially) that the types of variation which contribute to complex phenotypes may well follow much the same distribution as the types of mutations responsible for Mendelian diseases.

We now turn our attention to models of genetic variation, to describe how the basic genetic processes of mutation, transmission, recombination and selection act together to produce the patterns of variation observed.

Hardy–Weinberg and genetic drift

A ''locus'' is broadly defined in population genetics as a single genetic unit. It can be a single nucleotide site or a single contiguous region of a chromosome. The most basic model of polymorphism (genetic variation) in population genetics is a single locus with two alleles A and a, at frequencies p and q respectively, where $p+q=1$.

The first – and simplest – prediction of population genetics was described independently in 1908 by a British mathematician, G. H. Hardy, and a German obstetrician, Wilhelm Weinberg, and is appropriately named the Hardy–Weinberg Law (Hardy, 1908; Weinberg, 1908). This result predicts the relationship between allele frequencies and genotype

frequencies at a single, biallelic locus in an equilibrium (no drift, mutation, or selection) diploid population mating at random. In such a population, the Hardy–Weinberg Law states that the frequency of genotypes at an autosomal locus are simply the terms in the binomial expansion $(p+q)^2$. If p is the frequency associated with the A allele, then the terms p^2, $2pq$, and q^2 describe the frequencies of the AA, Aa, and aa genotypes respectively. This equilibrium can be thought of as the expectation of genotype frequencies under the random union of gametes and is reached after one generation of random mating. In practice, the prediction is robust to deviations from the strict assumptions listed above; so much so that departure from Hardy–Weinberg equilibrium is commonly used in human genetics as a method of detecting experimental errors (Gomes *et al.*, 1999; Xu *et al.*, 2002) or detecting disease susceptibility loci (Nielsen *et al.*, 1998; Wittke-Thompson *et al.*, 2005).

More generally, we would like to be able to model and predict the amount and distribution of genetic variation. It turns out that to do so, we need to understand the interplay between various forces; notably mutation, selection and genetic drift. It is therefore necessary to create a model that describes the change of our system of alleles over time. A very successful model for doing so, known as the Wright–Fisher model, was explored by Sewall Wright and R. A. Fisher during the 1920s. While this model is rather simple compared with the complicated realities of real populations, it does focus attention on the key concepts, and its predictions are often quite accurate in practice. The model is also readily modified to treat more realistic scenarios.

In the Wright–Fisher model, a population of N diploid individuals is treated as a collection of $2N$ haploid gametes. Reproduction of the population occurs simultaneously at discrete, nonoverlapping points in time. Each generation, the new population is made by sampling $2N$ gametes with replacement from the previous generation.

Consider the fate of an allele A which, at time t, is present in n_t copies in the population, and which therefore has a frequency $P_t = n_t/2N$. What is the frequency of A in the following generation? The Wright–Fisher model implies that the number of copies of the A allele in generation $t+1$ (call this n_t+1) is generated by binomial sampling of gametes during reproduction, so that in the absence of selection, $n_t+1 \sim$ Binomial $(P_t, 2N)$. That is,

$$\Pr(n_{t+1} = m) = \binom{2N}{m} P_t^m (1 - P_t)^{2N-m}, \tag{3.1}$$

so that

$$E(P_t + 1) = P_t \tag{3.2}$$

and

$$\mathrm{Var}(P_{t+1}) = \frac{P_t(1 - P_t)}{2N} \tag{3.3}$$

From Equation 3.2, we see that in the absence of natural selection, the average frequency of allele A is the same from one generation to the next. However, due to the randomness in the sampling process, the allele frequency will seldom be exactly the same in two consecutive generations (Figure 3.1). This is the essence of genetic drift. The variance (Equation (3.3)) measures the extent to which allele frequencies can differ between generations; this is a function of the present allele frequency and, importantly, the population size N. The rate of change of allele frequencies is inversely proportional to N, indicating that drift is accelerated in small populations.

If we start this process at time t, with an allele at frequency P_t, and then follow the frequency forward in time, the frequency drifts around over time. Eventually this process reaches one of two states: either $P=1$ (the A allele is fixed), or $P=0$ (the A allele is lost). Both of these states correspond to a loss of genetic variation. The allele frequencies then stay at 0 or 1 until there is a new mutation.

Thus, genetic drift tends to have the effect of reducing genetic variation, and this is a process that occurs faster in small populations than in large populations. At equilibrium, this effect is balanced by the input of new variation due to mutation.

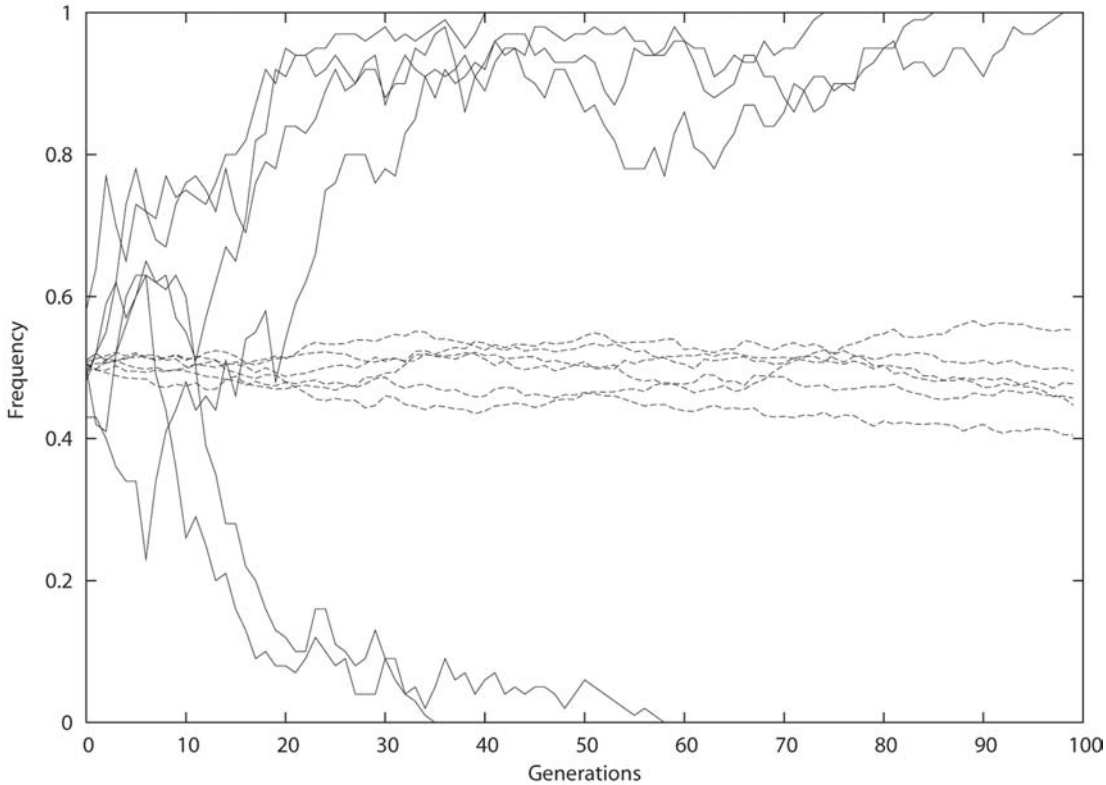

Figure 3.1 Genetic drift: the change in frequency of a neutral allele over time. Six simulations each of $t = 100$ generations are shown for $N = 100$ (solid lines) and $N = 10000$ (dashed lines), starting from an initial frequency of 0.5.

The heterozygosity of a locus is defined as the probability of obtaining different alleles in a random sample of two gametes from the population (or equivalently, the probability that an individual is heterozygous, assuming Hardy–Weinberg equilibrium), and can be written as

$$H = 1 - \sum_{i=1}^{n} p_i^2 \tag{3.4}$$

where p_i is the frequency of the *ith* allele. Although new alleles continue to emerge as the result of mutation, this is balanced by the loss of alleles due to drift, and at equilibrium the expected heterozygosity at a neutral locus is:

$$\mathrm{E}(H) = \frac{\theta}{\theta + 1} \tag{3.5}$$

where $\theta = 4N\mu$ (Crow and Kimura, 1970). This equilibrium state is referred to as mutation-drift balance. The parameter θ, also known as the population mutation parameter, plays a central role in population genetic theory. It contains all of the information necessary to predict the level of variation at neutral loci under the standard equilibrium model, and plays an important role in more general models as well. One important application of θ is the Ewens sampling theory (Ewens, 1972), which describes the number and frequencies of alleles expected under a neutral model (Hartl and Clark, 1997). For example, in a sample of size n, the expected number of alleles is given by

$$\mathrm{E}(K) = 1 + \frac{\theta}{\theta + 1} + \frac{\theta}{\theta + 2} + \cdots + \frac{\theta}{\theta + n - 1}. \tag{3.6}$$

The Ewens sampling theory predicts the distribution of allele frequencies at a neutral locus; this distribution, sometimes referred to as the site frequency spectrum, is the basis for many tests for selection used today (see below). The other major contribution of the Ewens sampling theory was to move the field of population genetics towards a sample-based method of inference. It was this movement that would ultimately give birth to the coalescent.

The coalescent

Thus far, we have described the dynamics of the Wright–Fisher model in terms of the random sampling of alleles, forwards in time. This was the most important analytical approach of classical population genetics. However, starting in the early 1980s, there has been an important conceptual shift towards thinking about the Wright–Fisher and related models in terms of the genealogical relationships between sampled chromosomes backwards in time (Kingman, 1982; Hudson, 1983; Tajima, 1983). The backwards process is referred to as the ''coalescent process'' (or simply ''the coalescent'').

It turns out that this alternative viewpoint is often preferable, as it can simplify mathematical analysis, and because it provides an extremely powerful tool for simulating data. Moreover, many questions of modern data analysis can be addressed much more easily within a coalescent framework than they can using the traditional forward modeling approach. Thus, while we do not have space to describe these models further here, we encourage the interested reader to consult reviews by Hudson (1990), Nordborg (2001), Rosenberg and Nordborg (2002) and the book by Hein *et al.* (2005).

Selection

The discussion so far has focused on neutral models of evolution. We now describe a simple model for natural selection. In this model, a zygote with genotype ij reaches reproductive age with some probability that is proportional to w_{ij}. The value w_{ij} is referred to as the ''fitness'' of genotype ij.

Then a model for the effect of a mutation on fitness is

Genotype	AA	Aa	aa
Frequency	p^2	$2pq$	q^2
Fitness	$w_{11}=1$	$w_{12}=1hs$	$w_{22}=1s$

Here s measures the strength of selection against homozygotes and h is the level of dominance of the a allele. When a is fully recessive, $h=0$; when a is fully dominant $h=1$. Overdominance (i.e. where the heterozygote has better fitness than either homozygote) can be modeled using $h>1$ and $s<0$. Within any given generation, the mean fitness of the population is

$$\bar{w} = p_2 w_{11} + 2pqw_{12} + q^2 w_{22} \tag{3.7}$$

and the change in frequency of the A allele from one generation to the next is described by,

$$p' = \left[\frac{p^2 w_{11} + pqw_{12}}{\bar{w}}\right](1 - p\mu_{Aa} + q\mu_{aA}), \tag{3.8}$$

where p' is the frequency of the A allele in the subsequent generation and the $(1-p\mu_{Aa}+q\mu_{Aa})$ term accounts for the net flux of alleles changed by mutation (here μ_{ij} is the mutation rate per generation from allele i to allele j).

Starting with this simple theory, there are several cases of particular interest. Firstly, and probably the most common type of selection in the human genome, is the case in which new mutations are disfavored. In that case, selection acts to remove the deleterious variants from the population; this is known as ''purifying'' or negative selection; this corresponds to $s<0$ in our model above (Charlesworth *et al.*, 1993). When there is strong purifying selection, as for most Mendelian diseases, the overall frequency of the disease allele in the

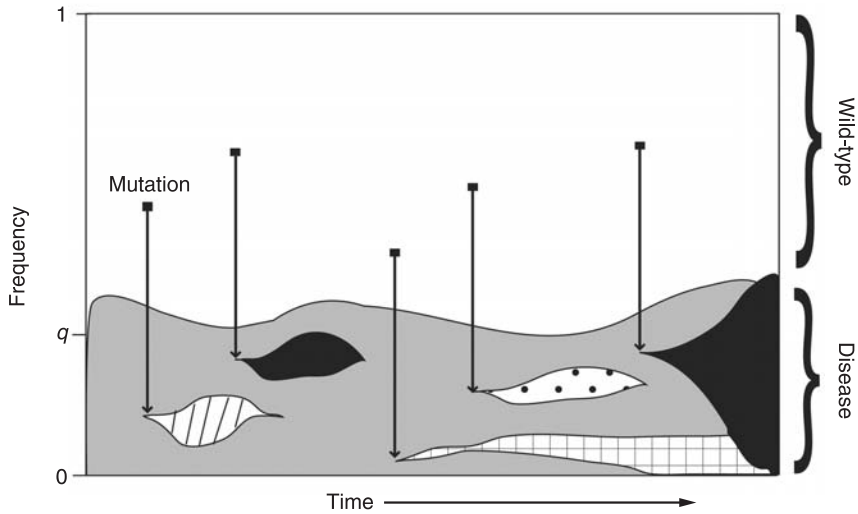

Figure 3.2 Mutation-selection balance. The degree of allelic heterogeneity at a Mendelian disease locus is a function of the mutation rate to the disease class θ_D, and the overall frequency of disease alleles *(q)* at that locus as determined by mutation-selection balance (Hartl and Campbell, 1982). Within the disease class, alleles will drift over time with new alleles replacing old ones. Creation of new disease alleles by mutation is indicated by an arrow leading from the wild-type allele population to the disease allele population.

population represents a balance between the input of new mutations and the removal of mutant alleles by natural selection. (Fig 3.2).

We can manipulate Equation (3.7) to predict the equilibrium allele frequencies by setting $p' = p$ and solving for the quantity q. It can be shown that for a fully recessive allele, the frequency at mutation-selection balance will be about

$$\hat{q} \approx \sqrt{\frac{\mu}{s}} \tag{3.9}$$

and for a partially dominant disease allele the equilibrium frequency is

$$\hat{q} \approx \frac{\mu}{hs} \tag{3.10}$$

where μ is the mutation rate at which new disease alleles arise from wildtype alleles. Several predictions can be made about mutations with strongly deleterious effects by analyzing Equations (3.9) and (3.10). First, the equilibrium allele frequency is predicted to be much lower for dominant diseases. This is because selection can ''see'' deleterious dominant alleles in heterozygous form and begin to eliminate them as soon as they arise. For recessive diseases, most copies of the disease allele are in ''carriers,'' i.e. heterozygous individuals who are asymptomatic.

A second important case occurs when a new mutation confers a reproductive advantage on its host over non-carriers, $s > 0$: such an allele is expected to increase in frequency in the population through natural selection. Population genetic theory predicts that the frequency of a strongly advantageous allele should increase quickly to fixation. There are various statistical tests that aim to detect signatures of positive selection left on the frequency distribution of neutral variation (Kreitman, 2000) or on linkage disequilibrium (see next section) surrounding a selected site (Slatkin and Bertorelle, 2001; Sabeti *et al.*, 2002; Voight *et al.*, 2006).

Lastly, when the heterozygous genotype is most fit ($h > 1$ and $s < 0$) or if selection acts in a frequency dependent manner, a stable

polymorphism at a locus can be maintained by balancing selection (e.g. G6PD and the MHC region; Tishkoff *et al.*, 2001; Garrigan and Hedrick, 2003).

The selective theory described above is deterministic, meaning that we ignore the role of random genetic drift. This is a good approximation when selection is strong (as a rough rule of thumb we might say that this is when $4Ns > 50$, say). It turns out that when $4Ns$ is not much greater than 1, the role of genetic drift becomes an important factor in determining the fate of alleles (e.g. Pritchard and Cox, 2002). For this reason, historically small populations, and populations which have experienced a bottleneck within the recent past may be enriched for disease mutations compared to our expectation from deterministic formulae (Risch *et al.*, 2003). In the weak selection case, we can model selection within the model for genetic drift described above by inserting p' in place of Pt in Equation (3.3).

Lastly, we note that there is a large and elegant literature on the theory of natural selection that we cannot do justice to in this limited space. The interested reader is encouraged to consult a text such as Hartl and Clark (1997) for further references.

Multi-locus models and linkage disequilibrium

Analysis of single locus models, as above, has yielded a great deal of insight into the dynamics of populations. However, genetic mapping methods rely fundamentally on the relationships between linked markers, both for linkage and association mapping. Therefore we must also understand the relationships between loci.

The simplest multi-locus model considers the relationship between two biallelic loci on the same haplotype, or gamete. Consider two loci, 1 and 2, segregating alleles '*A*' and '*a*', and '*B*' and '*b*', respectively. Then there are four allele frequencies: P_A, P_a, P_B and P_b, and four haplotype frequencies: P_{AB}, P_{Ab}, P_{aB} and P_{ab}. If an allele at site 1 is drawn at random from the population, it will be an '*A*' with probability p_A and an '*a*' with probability p_a; likewise with locus 2. Then by the rules of probability, if the allele frequencies at each locus are independent of one another, the probability of drawing a certain two-locus haplotype is just the product of the two individual allele frequencies. For example, assuming independence, the predicted frequency of the *AB* haplotype is simply P_AP_B. When two alleles are statistically independent of each other, they are said to be in linkage equilibrium. When individuals are sampled from a single population, it is normally the case that pairs of loci located on different chromosomes, or pairs that are far apart on the same chromosome, will be in linkage equilibrium. However, when loci are sufficiently close together, there is often statistical correlation between the alleles at each locus. This correlation is referred to as ''linkage disequilibrium'' or ''LD.''

Measures of LD

A large number of statistical measures have been proposed as ways of quantifying the level of LD between pairs of sites, or across regions of variation (Devlin and Risch, 1995; Hartl and Clark, 1997; Pritchard and Przeworski, 2001). None of these measures is ideal for all purposes; therefore we briefly describe three of the most prominent.

The most basic measure of LD between two sites is

$$D = P_{AB} - P_AP_B = P_{AB}P_{ab} - P_{Ab}P_{aB} \tag{3.11}$$

which takes a value of 0 when there is no LD (corresponding to independence between the sites).

Rewriting Equation (3.11) as $P_{AB} = P_AP_B + D$, we see that D may be interpreted as the amount by which a haplotype deviates from its equilibrium frequency. D has a more precise statistical interpretation as the covariance between allele frequencies at each site.

It is important to recognize that measurements of LD are sample-specific. Most depend on the

marginal allele frequencies (i.e. P_A and P_B) at each locus and the sample size of chromosomes used. The D' statistic, defined as D/D_{max} where

$$D_{\max} = \min\{P_A P_b, P_a P_B\} \qquad \text{when } D > 0 \qquad (3.12)$$

$$D_{\max} = \min\{P_A P_B, P_a P_b\} \qquad \text{when } D < 0 \qquad (3.13)$$

is a standardized statistic with range $-1 < D' < 1$ regardless of the allele frequencies in the sample. $|D'| = 1$ if and only if there are <4 haplotypes present in the sample, indicating no evidence of past recombination between the two sites.

Another key LD metric is r^2, the correlation coefficient between alleles at sites 1 and 2. It turns out that r^2 is the key determinant of power for association mapping if we are trying to map a disease mutation at an unknown position 2, using a marker at position 1 (Pritchard and Przeworski, 2001). r^2 is defined as

$$r^2 = \frac{D}{P_A P_B P_a P_b}. \qquad (3.14)$$

The value $n^* r^2$, where n is the sample size of chromosomes, corresponds to the standard chi-square test statistic for independence between alleles at sites 1 and 2. $r^2 = 1$ if there are only two 2-locus haplotypes present in the sample.

Recombination has the effect of reducing LD, so that in general, LD decays with the distance between two sites. At equilibrium in an infinite population, there would be no LD. However, in finite populations, genetic drift tends to move the haplotype frequencies away from the equilibrium state of no LD, so that in real populations there is a balance between drift and recombination, leading to stable LD at short distances. In a finite equilibrium population, the average level of LD between sites separated by a genetic distance c is

$$\mathrm{E}[r^2] = \frac{1}{1+\rho} \qquad (3.15)$$

where $\rho = 4Nc$ (reviewed by Hudson, 2001). In humans, detectable LD (i.e. departure from independence) typically exists between sites that are within about 10–100 kb apart, although instances of significant LD at much longer distances do occur (e.g. Reich *et al.*, 2001; Pritchard and Przeworski, 2001; Gabriel *et al.*, 2002). The structure and extent of LD in the human genome has important implications for the design and analysis of whole genome association mapping studies, such as the number and spacing of markers required to capture the majority of genetic variation in a population (Kruglyak, 1999; Carlson *et al.*, 2003). The goal of the ongoing HapMap project is to provide an empirical description of LD across the genome and should prove to be the sine qua non of aspiring gene-mappers (Altshuler *et al.*, 2005). As of January 2006, the HapMap consists of over 3.5 million SNPs genotyped in 270 individuals of African, East Asian, and European descent.

Models of complex disease

The common diseases which are the subject of this book are mainly complex traits that we cannot model with simple single-locus biallelic systems. We use the term ''genetic architecture'' to describe the biological basis for the translation of genetic variation into phenotypic variation. An excellent definition of genetic architecture may be found in Dilda and Mackay (2002): ''[a] list of all genes affecting variation in the trait; estimates of their additive, dominance, epistatic and pleiotropic effects and environmental sensitivities; and the molecular definition of [their] alleles''. To fully comprehend the genetic basis of a complex trait we must dissect and describe the genetic architecture of that trait.

However, in searching for complex trait loci, we are faced with the following dilemma: in order to describe the genetic architecture of a complex trait, we must be able to find the underlying genes. Experience indicates that this is difficult to do in humans, and theoretical results indicate that the choice of mapping strategy may have an important impact on the probability of success (e.g. Stephens *et al.*, 1994; Risch and Merikangas, 1996; Slager *et al.*, 2000). Unfortunately, the pros and cons of

the different strategies depend heavily on the unknown architecture of the trait!

Among the important and unknown aspects of genetic architecture that will impact study success are: the level of locus heterogeneity (i.e. the number of loci that contribute); the locus recurrence risk ratios (the size of effect due to each gene (Risch, 1990); the level of allelic heterogeneity (i.e. the number and frequencies of susceptibility alleles at each disease locus – important for LD-based mapping); the frequency distribution of risk alleles within populations; and the degree of differentiation across populations (for replication and for admixture mapping).

Common disease-common variant . . . or not?

One hypothesis that has become popular in recent years is that much of the genetic variation underlying complex traits is made up of common alleles (Lander, 1996; Chakravarti, 1999). This idea, often referred to as the ''common disease-common variant'' (CD-CV) hypothesis, has important implications for disease mapping, because it suggests that genome-wide association mapping should be an extremely powerful tool for identifying the variants that contribute to complex traits. An alternative model is that at most complex disease loci there are a few, or even many different alleles that all impact risk. This is the situation that is observed for most rare Mendelian traits. When the latter scenario occurs, the strength of association at single SNPs, or with haplotypes will usually be substantially weakened (Slager *et al.*, 2000), and so association mapping will be more difficult.

At the time of writing, there are not enough data to be able to evaluate the extent to which the CD-CV hypothesis is true. There are a number of examples of common variants for complex diseases that are known (Lohmueller *et al.*, 2003), and also clear examples of moderate allelic heterogeneity (at NOD2/CARD15, involved in Crohn's disease; Hugot *et al.*, 2001) and functional rare variants (associated with plasma levels of HDL cholesterol and LDL, and rates of sterol absorption; Cohen *et al.*, 2004; 2006). However, there is a clear ascertainment bias here. Relatively few complex trait loci have been identified thus far, and it is to be expected that many of the first loci to be discovered will be genetically simpler than the average.

Allelic heterogeneity at Mendelian loci

The loci that underlie Mendelian traits often exhibit extensive allelic heterogeneity (Terwilliger and Weiss, 1998; Estivill *et al.*, 1997). That is, there are often many different independent mutations at a single locus that cause disease. We now describe a simple population genetic model to understand allelic heterogeneity at such loci, and then briefly describe the extensions that have been used to predict what patterns of allelic heterogeneity are like at complex disease loci.

Consider a single locus at which there are two classes of alleles segregating at a given susceptibility locus – neutral and deleterious (Hartl and Campbell, 1982). If we make the assumption that within each class all alleles have the same effect on fitness then we can apply standard population genetic theory to make predictions about the number of alleles segregating at the locus. Let μ_D be the mutation rate at which disease alleles are formed per meiosis. Then the total frequency of new disease alleles each generation is

$$2N\mu_D(1-q) \approx 2N\mu_D. \tag{3.16}$$

If we assume that different disease mutations have the same selection coefficient, then

$$\theta_D = 4N(1-q)\mu_D \tag{3.17}$$

and the probability that two randomly chosen disease chromosomes have the same mutation at a given locus is about

$$\frac{1}{1+\theta_D}. \tag{3.18}$$

Thus, the degree of allelic heterogeneity depends mostly on the mutation rate to the disease class and not the overall frequency q of the disease class.

But in a growing population (like humans), θ is increasing, and so allelic heterogeneity should increase with time (Reich and Lander, 2001). Loci with low q should respond faster to the input of new mutation, and in fact they tend to have higher allelic heterogeneity when measured empirically.

Analytical results regarding allelic heterogeneity have only been obtained for the case of Mendelian disease loci, where we have the convenience of a simple genotype-to-phenotype relationship. An alternative approach to evaluating this issue for complex disease is to construct population genetic models (Pritchard, 2001; Reich and Lander, 2001). These models include mutation, natural selection and genetic drift, under some specified model of population history. While many of the important parameters are not completely known, we can at least use these models to get a sense of when CD-CV is likely to hold, and to what extent. These models indicate that the most important parameter determining the extent of allelic heterogeneity at a locus is the total mutation rate at which new risk alleles are generated at the locus. When this mutation rate is low, allelic heterogeneity will usually be low, but otherwise allelic heterogeneity may be moderate or high. Current guesses about the relevant mutation rate suggest that allelic heterogeneity may range from low to moderate at typical complex trait loci (reviewed by Pritchard and Cox, 2002).

Experimental results

To date, the best descriptions of genetic architecture for any complex trait have come from quantitative genetics experiments. Artificial selection experiments allow geneticists to enrich a study population for alleles with large effects on the phenotype of interest, and then design the most informative matings possible for trait mapping. This combination increases the power to detect and localize quantitative trait loci (QTLs) far beyond the capabilities of genetic experiments in humans. One of the most thorough of these experiments is a 2002 investigation of a classic complex trait – sensory bristle number in the fruit fly Drosophila melanogaster. This study by Dilda and Mackay (2002) used flies sampled from a large natural population and then selected for high- or low-bristle number over 25 generations of laboratory inbreeding. High- or low-bristle flies from these divergently selected laboratory populations were then crossed to highly inbred, wildtype (i.e. unselected) flies and the resulting recombinant inbred lines (RILs) were used in subsequent mapping for sensory bristle QTL. Two phenotypes were selected and scored in these crosses: abdominal and sternopleural bristle numbers.

The study identified 38 QTL for abdominal bristle number (AB) and 42 QTL for sternopleural bristle number (ST) after correcting for multiple comparisons. Along with the chromosomal location, the additive (main), dominance, and interaction effects of alleles segregating at each locus can be measured by QTL mapping. The distribution of additive effects of alleles followed a characteristic exponential or L-shaped distribution (Figure 3.3) that is predicted by theory (Robertson, 1967; Orr, 1998) and has been observed in the dissection of other quantitative traits (Hayes and Goddard, 2001) (Fig 3.3). This distribution corresponds to a few alleles of major effect and many alleles of minor effect. Dominance effects appeared to be randomly distributed around '0' (no dominance) with the majority of alleles showing an intermediate amount of dominance or recessivity; three alleles showed complete dominance, and one allele showed complete recessivity. More than 30% of the total genetic variance in bristle number, pooled across all crosses, was due to interaction effects. These interactions can be partitioned into sex specific effects, gene–environment and gene–gene interactions.

Sex Effects

All 38 QTL mapped for AB showed sex-specific effects. Of these, 22 were conditionally neutral: loci

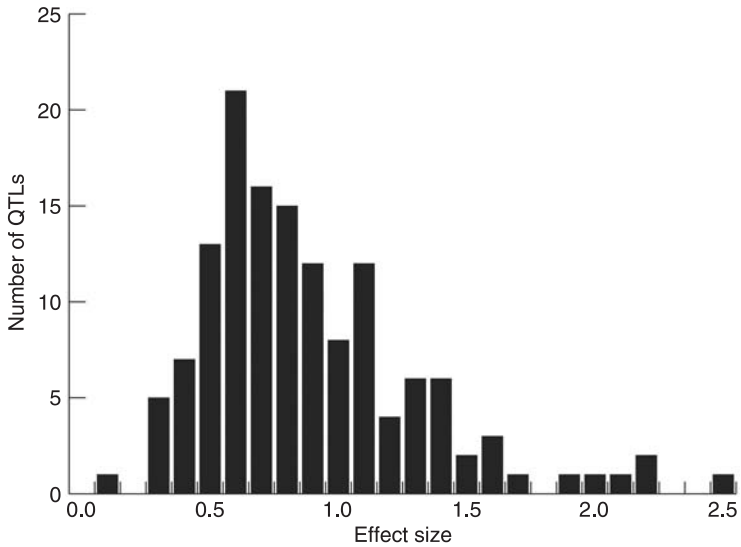

Figure 3.3 L-shaped distribution of additive QTL effects from (Dilda and Mackay, 2002). Note the threshold of observed QTL near 0.4. Theory predicts that there are many more QTL of effect <0.4 that were not detected due to sample size limitations. Additive effects are given in standard deviations (a/σ_a).

that showed significant association with bristle number in one gender but not the other. The remaining 16 loci mapped to AB displayed opposite effects between males and females, a phenomenon known as antagonistic pleiotropy. Among ST lines 19% (8/42) loci showed sex-specific effects, all of which were conditionally neutral.

Gene (x) environment interaction (GEI)

Three separate groups of progeny were selected from each cross and reared at either 18°C, 25°C, or 28°C. QTL mapping identified many loci that showed temperature-dependent effects for each phenotype. 33% (14/42) of ST QTL and 55% (21/38) of AB fell into this category. The more remarkable observation was that 92% (20/21) of AB and 78.6% (11/14) of ST loci involved in GEIs were only detected at a single temperature. The temperatures used here were arbitrary, discrete categories that imperfectly capture the continuous variation of environment in nature.

Epistasis

A pairwise search of all QTL displaying main effects identified 13 statistically significant gene–gene interactions with effect on AB and 26 interactions for ST loci. 8/13 AB epistatic interactions were due to the greater-than-additive (synergistic) effect of alleles, while 22/26 ST interactions were less-than-additive (''diminishing'' epistasis).

Summary

Interaction effects are extremely important because they cause the additive effects of alleles to change as the genetic and environmental background of a population changes. Mapping studies are often best at detecting additive effects; therefore the prevalence of GEI and epistasis that is emerging from model organism research has profound implications for the success and reproducibility of disease mapping studies in humans. A large number of QTL identified by Dilda and

Mackay (2002) were only identified within a single sex or environment. As an accurate mathematical parameterization of the complex and continuously changing human environment may be impossible, only alleles with strong effects, or effects independent of environment, will be detected in most studies of cosmopolitan populations. Likewise, we should not be surprised when associations identified in one population are not reproduced in another (Wade, 2001). The complexity of metabolic networks suggested by our limited number of coding sequences, coupled with the great diversity of our geographic and social climates, indicates the potential for a greater number of interactions in man than in any other organism. Finally, we note that the genetic architecture of *Drosophila* bristle number, as it appears in a selection experiment, may be qualitatively different from the genetic architecture of a complex disease phenotype in humans. The degree of disparity will depend partly on the nature of selection acting on the disease phenotype, which may be limited due to the post-reproductive onset of many common multifactorial diseases (Wright *et al.*, 2003).

REFERENCES

Altshuler, D., Brooks, L. D., Chakravarti, A. *et al.*; International HapMap Consortium. (2005). A haplotype map of the human genome. *Nature*, **437**, 1299–320.

Baird, P. A., Anderson, T. W., Newcombe, H. B. and Lowry, R. B. (1988). Genetic disorders in children and young adults: a population study. *Am J Hum Genet*, **42**, 677–93.

Botstein, D. and Risch, N. (2003). Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat Genet*, **33** Suppl, 228–37.

Carlson, C. S., Eberle, M. A., Rieder, M. J. *et al.* (2003). Additional SNPs and linkage-disequilibrium analyses are necessary for whole-genome association studies in humans. *Nat Genet*, **33**, 518–21.

Chakravarti, A. (1999). Population genetics–making sense out of sequence. *Nat Genet*, **21**, 56–60.

Charlesworth, B., Morgan, M. T. and Charlesworth, D. (1993). The effects of deleterious mutations on neutral variation. *Genetics*, **134**, 1289–303.

Cohen, J. C., Pertsemlidis, A., Fahmi, S. *et al.* (2006). Multiple rare variants in NPC1L1 associated with reduced sterol absorption and plasma low-density lipoprotein levels. *Proc Natl Acad Sci USA*, **103**, 1810–15.

Cohen, J. C., Kiss, R. S., Pertsemlidis, A. *et al.* (2004). Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science*, **305**, 869–72.

Conrad, D. F., Andrews, T. D., Carter, N. P., Hurles, M. E. and Pritchard, J. K. (2006). A high-resolution survey of deletion polymorphism in the human genome. *Nat Genet*, **38**, 75–81.

Crow, J. F. (2000). The origins, patterns and implications of human spontaneous mutation. *Nat Rev Genet*, **1**, 40–7.

Crow, J. F. and Kimura, M. (1970). *An introduction to population genetics theory*. New York: Harper and Row.

Devlin, B. and Risch, N. (1995). A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics*, **29**, 311–22.

Dilda, C. L. and Mackay, T. F. (2002). The genetic architecture of Drosophila sensory bristle number. *Genetics*, **162**, 1655–74.

Estivill, X., Bancells, C. and Ramos, C. (1997). Mutations in European populations. The Biomed CF Mutation Analysis Consortium. *Hum Mutat*, **10**, 135–54.

Ewens, W. J. (1972). The sampling theory of selectively neutral alleles. *Theor Pop Gen*, **3**, 87–112.

Falconer, D. S. and Mackay, T. F. C. (1996). *Introduction to quantitative genetics*. Essex, UK: Longman Group Ltd.

Fay, J. C., Wyckoff, G. J. and Wu, C. I. (2001). Positive and negative selection on the human genome. *Genetics*, **158**, 1227–34.

Gabriel, S. B., Schaffner, S. F., Nguyen, H. *et al.* (2002). The structure of haplotype blocks in the human genome. *Science*, **296**, 2225–9.

Garrigan, D. and Hedrick, P. W. (2003). Perspective: detecting adaptive molecular polymorphism: lessons from the MHC. *Evolution*, **57**, 1707–22.

Gillespie, J. (2004). *Population genetics, a concise guide*, 2nd edn. Baltimore, MD: Johns Hopkins University Press.

Gomes, I., Collins, A., Lonjou, C. *et al.* (1999). Hardy-Weinberg quality control. *Ann Hum Genet*, **63**, 535–8.

Haldane, J. B. S. (1947). The mutation rate of the gene for hemophelia and its segregation ratios in males and females. *Ann Eugen*, **13**, 262–71.

Hardy, G. H. (1908). Mendelian proportions in a mixed population. *Science*, **28**, 49–50.

Hartl, D. and Clark, A. G. (1997). *Principles of population genetics*, 3rd edition Sunderland, MA: Sinauer Associates Inc.

Hartl, D. L. and Campbell, R. B. (1982). Allele multiplicity in simple Mendelian disorders. *Am J Hum Genet*, **34**, 866–73.

Hayes, B. and Goddard, M. E. (2001). The distribution of the effects of genes affecting quantitative traits in livestock. *Genet Sel Evol*, **33**, 209–29.

Hein, J., Schierup, M. K. and Wiuf, C. (2005). *Gene genealogies, variation and evolution: a primer in coalescent theory*. New York, USA: Oxford University Press.

Hudson, R. R. (1983). Properties of a neutral allele model with intragenic recombination. *Theor Popul Biol*, **23**, 183–201.

Hudson, R. R. (1990). *Gene genealogies and the coalescent process.*, volume 7. Oxford, UK: Oxford University Press, pp. 1–43.

Hudson, R. R. (2001). Linkage disequilibrium and recombination. In Balding, D., Bishop, M. and Cannings, C. eds., *Handbook of statistical genetics*. Wiley, pp. 309–24.

Hugot, J. P., Chamaillard, M., Zouali, H. *et al.* (2001). Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature*, **411**, 599–603.

Iafrate, A. J., Feuk, L., Rivera, M. N. *et al.* (2004). Detection of large-scale variation in the human genome. *Nat Genet*, **36**, 949–51.

Kingman, J. F. C. (1982). On the genealogy of large populations. *J Appl Probab*, **19A**, 27–43.

Kondrashov, A. S. (1995). Contamination of the genomes by very slightly deleterious mutations. why have we not died 100 times over? *J Theor Biol*, **175**, 583–94.

Kondrashov, A. S. (2003). Direct estimates of human per nucleotide mutation rates at 20 loci causing Mendelian diseases. *Hum Mutat*, **21**, 12–27.

Kreitman, M. (2000). Methods to detect selection in populations with applications to the human. *Ann Rev Genomics Hum Genet*, **1**, 539–59.

Kruglyak, L. (1999). Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat Genet*, **22**, 139–44.

Lander, E. S. (1996). The new genomics: global views of biology. *Science*, **274**, 536–9.

Lohmueller, K. E., Pearce, C. L., Pike, M., Lander, E. S. and Hirschhorn, J. N. (2003). Metaanalysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet*, **33**, 177–82.

Nachman, M. W. and Crowell, S. L. (2000). Estimate of the mutation rate per nucleotide in humans. *Genetics*, **156**, 297–304.

Nielsen, D. M., Ehm, M. G. and Weir, B. S. (1998). Detecting marker-disease association by testing for Hardy-Weinberg disequilibrium at a marker locus. *Am J Hum Genet*, **63**, 1531–40.

Nordborg, M. (2001). Coalescent theory. In Balding, D., Bishop, M. and Cannings, C. eds., *Handbook of statistical genetics*. Chichester, England: Wiley, pp. 179–212.

Orr, H. A. (1998). The population genetics of adaptation: the distribution of factors fixed during evolution. *Evolution*, **52**, 935–49.

Pritchard, J. K. (2001). Are rare variants responsible for susceptibility to complex diseases? *Am J Hum Genet*, **69**, 124–37.

Pritchard, J. K. and Cox, N. J. (2002). The allelic architecture of human disease genes: common disease-common variant . . . or not? *Hum Mol Genet*, **11**, 2417–23.

Pritchard, J. K. and Przeworski, M. (2001). Linkage disequilibrium in humans: models and data. *Am J Hum Genet*, **69**, 1–14.

Reich, D. E., Cargill, M., Bolk, S. *et al.* (2001). Linkage disequilibrium in the human genome. *Nature*, **411**, 199–204.

Reich, D. E. and Lander, E. S. (2001). On the allelic spectrum of human disease. *Trends Genet*, **17**, 502–10.

Rimoin, D. L., Connor, J. M. and Pyeritz, R. E. (1997). *Nature and frequency of genetic disease*. 3rd edn. Edinburgh: Churchill Livingstone, pp. 31–4.

Risch, N. (1990). Linkage strategies for genetically complex traits I. Multilocus models. *Am J Hum Genet*, **46**, 222–8.

Risch, N. and Merikangas, K. (1996). The future of genetic studies of complex human diseases. *Science*, **273**, 1516–17.

Risch, N., Tang, H., Katzenstein, H. and Eckstein, J. (2003). Geographic distribution of disease mutations in the Ashkenazi Jewish population supports genetic drift over selection. *Am J Hum Genet*, **72**, 812–22.

Robertson, A. (1967). The nature of quantitative genetic variation. In Brink, A. ed., *Heritage from Mendel*. Madison, WI: The University of Wisconsin Press, pp. 265–80.

Rosenberg, N. A. and Nordborg, M. (2002). Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. *Nat Rev Genet*, **3**, 380–90.

Sabeti, P. C., Reich, D. E., Higgins, J. M. *et al.* (2002). Detecting recent positive selection in the human genome from haplotype structure. *Nature*, **419**, 832–7.

Sachidanandam, R., Weissman, D., Schmidt, S. C. *et al.* (2001). A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature*, **409**, 928–33.

Salisbury, B. A., Pungliya, M., Choi, J. Y. *et al.* (2003). SNP and haplotype variation in the human genome. *Mutat Res*, **526**, 53–61.

Sebat, J., Lakshmi, B., Troge, J. *et al.* (2004). Large-scale copy number polymorphism in the human genome. *Science*, **305**, 525–8.

Slager, S. L., Huang, J. and Vieland, V. J. (2000). Effect of allelic heterogeneity on the power of the transmission disequilibrium test. *Genet Epidemiol*, **18**, 143–56.

Slatkin, M. and Bertorelle, G. (2001). The use of intraallelic variability for testing neutrality and estimating population growth rate. *Genetics*, **158**, 865–74.

Stenson, P. D., Ball, E. V., Mort, M. *et al.* (2003). Human gene mutation database (HGMD): 2003 update. *Hum Mutat*, **21**, 577–81.

Stephens, J. C., Briscoe, D. and O'Brien, S. J. (1994). Mapping by admixture disequilibrium in human populations: limits and guidelines. *Am J Hum Genet*, **55**, 809–24.

Tajima, F. (1983). Evolutionary relationships of DNA sequences in finite populations. *Genetics*, **105**, 437–60.

Terwilliger, J. D. and Weiss, K. M. (1998). Linkage disequilibrium mapping of complex disease: fantasy or reality? *Curr Opin Biotechnol*, **9**, 578–94.

Tishkoff, S. A., Varkonyi, R., Cahinhinan, N. *et al.* (2001). Haplotype diversity and linkage disequilibrium at human G6PD: recent origin of alleles that confer malarial resistance. *Science*, **293**, 455–62.

Tuzun, E., Sharp, A. J., Bailey, J. A. *et al.* (2005). Fine-scale structural variation of the human genome. *Nat Genet*, **37**, 727–32.

Voight, B. F., Kudaravalli, S., Wen, X. and Pritchard, J. K. (2006). A map of recent positive selection in the human genome. *PLoS Biol*, **4**, e72.

Wade, M. J. (2001). Epistasis, complex traits, and mapping genes. *Genetica*, **112**, 59–69.

Weinberg, W. (1908). Ueber den Nachweis der Vererbung beim Menschen. *Jh Ver Vaterl Naturk Wurttemb*, **64**, 368–82.

Wittke-Thompson, J. K., Pluzhnikov, A. and Cox, N. J. (2005). Rational inferences about departures from Hardy-Weinberg equilibrium. *Am J Hum Genet*, **76**, 967–86.

Wright, A., Charlesworth, B., Rudan, I., Carothers, A. and Campbell, H. (2003). A polygenic basis for late-onset disease. *Trends Genet*, **19**, 97–106.

Xu, J., Turner, A., Little, J., Bleecker, E. R. and Meyers, D. A. (2002). Positive results in association studies are associated with departure from Hardy-Weinberg equilibrium: hint for genotyping error? *Hum Genet*, **111**, 573–4.

4

Mapping common disease genes

Naomi R. Wray and Peter M. Visscher

Introduction

What is ''gene mapping'' and why is it useful?

One of the goals of human genetics research is to understand genetic variation between people in their susceptibility to disease. From twin and family studies, and the study of Mendelian disease, it is clear that some traits and diseases ''run in families'' and that the reason for the increased disease risk of relatives of affected individuals is, at least in part, because of their genetic predisposition. Genetic variation in populations is caused by mutations that cause differences in DNA sequence and by other genome events in the germline, for example insertions, deletions, duplications and translocations of stretches of DNA. If these mutation events have an effect on a phenotype of the carrier, for example an increased risk of disease or an effect on a continuously varying phenotype (such as blood pressure or body mass index), then there will be an association between the genotype and the phenotype. Gene mapping aims to identify locations on the genome that are responsible for genetic variation and, ultimately, to identify which specific variants cause the observed effect. Gene mapping is useful because it leads to an understanding of the nature of genetic variation and the identification of variants and biological pathways that cause or predispose to disease. This knowledge can be used to develop drug targets or other treatments and in the future may be used for disease diagnosis or the assessment of susceptibility to disease. The knowledge resulting from gene mapping is also useful in evolutionary studies and in plant and animal breeding.

Variation and disease

The most common type of genetic variation in the human genome is single DNA base changes, called single-nucleotide-polymorphisms or SNPs. There are 10 to 15 million ''common'' SNPs in humans, where ''common'' is usually defined as a frequency of at least 1% of the minor allele (that is, the allele with the smallest frequency). Because germline mutation events at nucleotides are, on average, rare (of the order of 1 in 100 million per site per generation), most SNPs only have two alleles at any one time in the population. In theory it is possible to have four alleles, corresponding to each of the four nucleotide bases (A, C, G and T).

Even with modern highly automated platforms for genotyping multiple genetic markers, usually individuals are genotyped for only a small fraction of all the polymorphic sites in the genome. For example, the currently available high-density 300 000 and 500 000 SNP arrays for humans contain approximately 3–5% of all common polymorphisms, and a much smaller proportion of rare variants. Fortunately, information about all polymorphic sites from a sample can be inferred. An important reason why this inference is possible is linkage disequilibrium (LD), i.e. the non-random association between alleles at different loci. LD equates to a correlation between the presence of

specific alleles at different loci on the same chromosome. The correlation structure between multiple markers in a region of the genome allows only a subset of markers to be genotyped, so that unobserved genotypes are correlated to the ones that are genotyped (see Chapter 3). LD is caused by mutation, genetic drift, inbreeding, founder effects, population stratification and selection. It is broken down by recombination (Hartl and Clark, 1997).

If the mutation rate per nucleotide per generation is 2×10^{-8} and there are approximately 3×10^{9} base pairs in the genome, then on average every individual receives about 60 new mutations from each parent. The majority of these (very rare) mutations are likely to be neutral, i.e. they do not have an adverse effect on fitness, including predisposition to disease, but a few may have deleterious effects. The impact on the disease burden in a population is a function of the frequency of the variant in the population and its effect on disease. Most deleterious variants will not have a significant effect on disease burden because either their frequency is too small and/or their effect is small. It is for this reason that current endeavors in human complex disease genetics are focused on common variants with a detectable effect.

Mendelian disease is characterized by a strong, usually one-to-one, relationship between the genotype of an individual and his/her phenotype. For example, in cystic fibrosis, two susceptibility alleles are necessary and sufficient to cause disease, whereas in Huntington's disease a single (dominant) mutation is responsible for the disorder. Common ''complex'' human disease, for example cardiovascular disease, asthma and psychiatric disorders, is characterized by a weaker relationship between genotype and phenotype. Monozygotic twins can be discordant for such disorders: individuals carrying susceptibility alleles may not become ill and others may become ill, whilst not having specific susceptibility alleles. This uncertainty is usually quantified in terms of probabilities or ratios. For example, a relative risk of disease to relatives, a probability of disease given the presence of a susceptibility genotype (penetrance) and the probability of disease given the absence of a susceptibility genotype (phenocopy rate). Mendelian disease loci are relatively easy to map because of the strong phenotype–genotype relationship, whilst complex disease loci have an inbuilt uncertainty.

Principles of gene mapping

The aim of gene mapping is to position genetic variants on the genome, relative to observed genetic variants with a known location. Usually the genes or loci to be mapped are those that predispose to disease, but gene mapping can also apply to new marker loci or to loci that cause variation in continuously varying phenotypes. There are two main approaches to gene mapping which are discussed in this chapter: linkage mapping in pedigrees and linkage disequilibrium mapping in the population (association studies). These approaches have similarities because pedigrees are a set of related individuals which are a subset from the whole population, but differ greatly in experimental design, statistical power, map resolution and the risk of false positives.

Linkage mapping in families

The principle of linkage mapping in pedigrees is the co-segregation of a phenotype and a genetic marker within families. Markers and disease loci that are closely linked on the same chromosome will tend to be inherited together, so that affected offspring tend to share the same chromosome segments (because there are only a few recombination events on a chromosome in each meiosis). The essence of most statistical tests for linkage is the counting of recombination events: if a susceptibility locus is unlinked to a marker, then there should be no relationship between phenotype and marker genotype, and the observed recombination rate between marker and phenotype should not be significantly different from 0.5, the value for free recombination.

The statistical power to detect linkage in a pedigree depends on the ascertainment of the pedigrees, the informativeness of the genetic markers, the relationship between genotype and phenotype and the sample size. Usually pedigrees are ascertained (selected) because multiple individuals are affected. This ascertainment scheme is employed to enrich the sample with disease variants so that power to detect disease loci is increased. An extreme example is the ascertainment of families in which rare Mendelian disease alleles are segregating. Informativeness of genetic markers can be measured by heterozygosity, i.e. the probability that two alleles at the locus are different. Microsatellite markers have been the loci of choice for pedigree linkage studies, because of their abundance in the genome (10 000s) and their high level of heterozygosity (typically 70 to 80%). High heterozygosity implies that individuals are likely to have two different alleles at a marker locus and that other individuals are likely to have different alleles. This makes the tracking of marker alleles through a pedigree easier and minimises ambiguity in estimating which DNA segments are shared identical-by-descent between relatives. Apart from the selection of families enriched for segregating disease alleles, the relationship between genotype and phenotype cannot be controlled experimentally but a larger sample size means that loci of smaller effects can be detected. For linkage studies, a usual measure of sample size is the number of informative meioses in the pedigree, i.e. the number of chromosomes that can be tracked from parents to progeny in the entire pedigree.

Experimental designs

In human populations a number of different experimental designs have been employed to map susceptibility loci for common disease in families. The choice is governed by the availability and ease of sample collection and statistical power. Sampling affected pairs of relatives, in particular pairs of affected siblings, is a popular design because phenotypic and genotypic information on other relatives is not necessary, making it particularly useful for late-onset disease, and because many such pairs exist if the disease is common. For example, if the population incidence of a disorder is 0.10 (e.g. recurrent major depression) and the relative sibling risk is 2.0, then the probability of observing a pair of affected siblings is $2 \times 0.1 \times 0.1 = 0.02$. Hence, 2% of all sibling pairs in the population are affected. In an affected sibling design, the sibs are genotyped for markers throughout the genome, typically 400 to 800 microsatellite markers, and the test statistic for linkage is based upon the proportion of alleles that the sibs share identical-by-descent (IBD). On average, sibs share 50% of their genes IBD at a particular location in the genome. If they are affected because of a variant that they both inherited from one or both parents, then the proportion of alleles shared IBD at and around the susceptibility locus will be more than 50%. In order to detect linkage it is not necessary that the same disease alleles segregate across all families. If there are a number of different rare variants at one or multiple disease susceptibility loci in the same genome region, then linkage can detect this region because excess IBD sharing is not with respect to a particular allele in the population. This is important, and contrasts with the population-based association studies which are discussed subsequently. Figure 4.1 shows the power to detect linkage in a genome scan for affected pairs of siblings, as a function of the relative risk of a carrier of the susceptibility allele. It was assumed that there are only two alleles at the disease locus in the population and that the risk is multiplicative, so that if the relative risk of a heterozygote is 3 then an individual with two copies of the susceptibility allele has a nine-fold risk of becoming ill relative to the population risk. The frequency of the susceptibility allele in the population was assumed to be 0.10 and the statistical test of significance was the mean proportion of alleles shared IBD against its expectation of 0.5 (Sham, 1998). It can be seen that if the relative risk of a heterozygote is more than three-fold

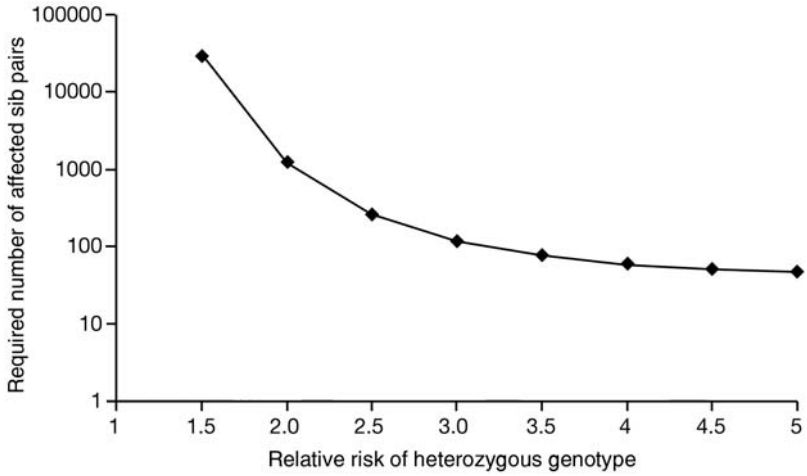

Figure 4.1 Number of affected sibpairs to detect linkage to a disease locus with a frequency of the susceptibility allele of 0.10 and a multiplicative relative risk of genotypes, for a power of 80% and a type I error rate of 0.0001.

then fewer than 100 affected pairs are required to reach genome-wide significance. However, many more pairs are needed for more realistic relative risks at susceptibility loci. For relative risks less than 1.5, 1000s of affected sibling pairs are needed, which becomes very difficult and expensive to ascertain in the population. As will be seen, population-wide association studies are more powerful for disease loci with small relative risks.

Selection of nuclear families or extended ''complex'' pedigrees are also ascertained with respect to the number of affected individuals, to enrich the sample for segregating disease alleles. It is hoped that large multiple generation pedigrees which are ascertained for the prevalence of disease are more homogeneous with respect to the number of segregating disease alleles than selections of small families for the same disease. However, there is insufficient evidence to show that this is generally true for common disease. Within a population, all individuals will be related if we could trace the pedigree back far enough, so that the distinction between ''unrelated'' small families and large extended pedigrees becomes blurred at some point. If alleles for common disease are common then many families will be segregating for these alleles. Nevertheless, large pedigrees generally have an advantage of statistical power, because many contrasts between relatives in the pedigree (e.g. sib pairs, child–grandparent, cousins) are informative for linkage. In extended pedigrees, the information on linkage decreases with a decrease in the genetic relationship between pairs of relatives. In large pedigrees, most linkage information is usually provided by sibling pairs.

The occurrence of multiple cases of rare disorders in families where parents are related, i.e. from consanguineous marriages is suggestive of a recessive genetic model. These families provide a special opportunity to map genes for rare recessive disorders. If an allele that predisposes to disease acts recessively, i.e. two copies of the risk allele are needed to become affected, and is at a low frequency in the population, then there will be relatively few homozygous individuals in the population. If the genotypes are in Hardy–Weinberg equilibrium (see Chapter 3) and the risk allele frequency is p, then the proportion of homozygotes and carriers is p^2 and $2p(1\text{-}p)$, respectively. Hence, for $p = 0.001$, 1 in a million individuals have two copies and are affected, but 1 in 500 individuals are carriers. The reason why ascertaining affected

offspring from consanguineous marriages is a powerful design for recessive diseases is that the parents have both been enriched to be carriers of the same risk allele that they inherited from a recent ancestor. The probability that an individual has inherited two risk alleles given that his parents are related is approximately pF, with F the inbreeding coefficient of the individual. For $p = 0.001$ and $F = 1/16$ (e.g. a child whose parents are first cousins), this means that the probability is 1 in 16 000, much higher than the probability of 1 in a million for non-inbred individuals. The implication of this is that an affected offspring of a consanguineous marriage will be homozygous by descent at and around the locus that predisposes to disease, because the exact same segment of DNA has been inherited from a common ancestor. For example, if the parents are first cousins then the affected child is likely to have inherited two copies of the same risk allele from one of the great-grandparents. The gene mapping method for such ascertained cases is called homozygosity mapping. Affected individuals are genotyped and the aim is to find regions of the genome where all or most cases are homozygous for a segment of DNA. The same method can be used with distantly related affected individuals even if their complete genealogy is unknown, for example, if they are all from an isolated community. Affected individuals may have inherited two copies of the same risk allele from a more distant relative. Homozygosity mapping can be successful with only a few cases. For example, a locus for benign recurrent intrahepatic cholestasis, a rare recessive disease, was mapped to chromosome 18 using only three affected individuals (Houwen *et al.*, 1994).

A linkage study may result in the detection of a significant linkage region of the order of 10–30 Mb. This is an important improvement in scale for the pursuit of gene mapping compared to the total genome length of 3000 Mb, but nonetheless can represent a region of many hundreds of known as well as unknown genes. There are approximately 20 000 genes in the human genome, so on average 7 per million base pairs (with a lot of variation around this average because some regions in the genome are gene-poor and others gene-rich). Hence a typical linkage regions contains 100 to 200 genes. The length of the linkage region is a direct reflection of the small number of meioses and the long length of shared DNA between affected family members. The underlying causal variants of different families that show evidence for linkage may or may not be the same.

There are a large number of different statistical tests used in practice to test for linkage, depending on the design, the amount of marker information and the phenotype (discrete or continuous) (Sham, 1998). Available software packages for linkage analysis can be found at linkage.rockefeller.edu/soft/list.html. In human genetics, the strength of evidence for linkage is often expressed as a LOD (Likelihood of Odds) score, or simply as a p-value, the probability of finding a more extreme value of the test statistic by chance. The LOD score is the logarithm to the base 10 of the ratio of the likelihoods for linkage (the alternative hypothesis) and for no linkage (the null hypothesis). For genome scans, LOD values of approximately 2.0 and 3.0 correspond to ''suggestive'' and ''significant'' linkage (Lander and Kruglyak, 1995). Suggestive linkage means that the LOD score is expected to exceed the value of 2.0 by chance (i.e. under the null hypothesis) once per genome scan. Significant linkage means that the probability of exceeding the threshold of 3.0 is approximately 0.05, i.e. the LOD score is expected to exceed this value by chance once in every 20 genome scans. The LOD score is often used because of its intuitive interpretation, i.e. a 10-fold difference in the ratio of the likelihoods per unit. The 'magic' threshold of LOD = 3, equivalent to a ratio of the likelihoods of 1000 was first suggested by N. Morton (Morton, 1955) in the context of linkage of a known Mendelian trait to markers in the genome. Morton assumed type-I and type-II error rates of 0.001 and 0.01, respectively, and a prior probability of linkage of 0.05 to derive a 'posterior probability of a type-I error' of 0.02, using the theory of sequential testing. Study-specific LOD thresholds

can be obtained from computer simulations or from a permutation analysis. The advantage of a study-specific LOD threshold for significance is that it does not rely on assumptions from which the guideline thresholds are calculated, and properly takes the sample size and marker informativeness of the sample into account.

Association studies

In contrast to linkage studies, association studies look for a correlation between a specific variant and disease status or quantitative trait in the population. Traditionally, association studies were set up for candidate genes, which were either positional candidates located within a region identified by a linkage study or functional candidates whose known function seemed of relevance to the disease of interest. With the advent of high-throughput genotyping of SNPs, association scans of complete linkage regions or whole-genome scans, requiring the genotyping of 100 000s of loci are now possible.

The success of association studies depends on the genetic architecture underlying complex diseases. The Common Disease Common Variant (CDCV) hypothesis (Chakravarti, 1999) predicts that disease risk alleles underlying complex diseases will be relatively common and this is the foundation stone of current association studies. The usual assumption has been that disease variants arose in a single common ancestor many generations ago and that the variants are now at frequency of $>1\%$ in the population as it is unlikely that exactly the same variants have arisen multiple times. More recently, the ancestral-susceptibility model has been proposed (Di Rienzo and Hudson, 2005). This model still assumes that common variants will be responsible for common diseases, but proposes that the ancestral allele at a disease risk locus is the susceptibility variant and that new mutations at the locus are protective. Empirical evidence for ancestral susceptibility variants can be found in a number of ''modern diseases'' such as asthma, obesity, type 2 diabetes, coronary heart disease and Alzheimer's disease whose frequency has increased in recent generations because of changes in lifestyle and life expectancy. In evolutionary history, susceptibility variants for these diseases may not have been under selection pressure and indeed may have conferred selection advantage. For example, risk variants for type 2 diabetes expressed in our nutritionally enriched environment are likely to have conveyed selection advantage in a nutritionally limiting environment (Di Rienzo and Hudson, 2005). We are now entering an era of genome-wide association studies in which several hundred thousand SNPs are genotyped over the entire genome. Results from these studies should provide insight into the true genetic architecture of complex diseases.

A comprehensive association study of a chromosomal region requires a dense spacing of markers. SNPs are the marker of choice in association studies because they occur on average more than once each 1000 base pairs, they are amenable to high-throughput scoring methods and they are relatively stable over the generations of meioses from the common ancestor. It is estimated that there are up to 15 million common SNPs in the human genome. It is not necessary to genotype them all as there is redundancy in the information between syntenic (located on the same chromosome) SNPs because of the linkage disequilibrium between them. Association of a disease with a genotyped variant may reflect association with a causal variant that is in LD with the genotyped SNPs.

Once an association is found between a genetic variant and disease, it may take many years and many research dollars to identify the causal functional variant(s). Following up on false positives can be avoided by considering only replicated association results which is particularly important given the level of multiple testing which is characteristic of association studies. Association is a statistical rather than causal concept and other factors may lead to detection of association which are not the result of an association with a causal

genetic variant. For example, association may be detected if the population study contains genetically distinct subgroups, so-called ''population stratification''. If both the disease and a genetic variant are more prevalent in one subgroup then a spurious association between the disease and variant will be detected in the entire population. Spurious association may also be detected for late-onset diseases because of natural selection, whereby the associated variant is causal with respect to the ability to survive and, by association, to express the disease.

The stages in conducting an association study are (1) power or sample size calculations; (2) collection of samples and DNA; (3) SNP selection from the chromosomal region(s) under study; (4) genotyping of the study sample; (5) analysis and interpretation. Each of these stages will be considered in turn, but first measurement of linkage disequilibrium is introduced as understanding the concept of linkage disequilibrium is fundamental to understanding the design and analysis of candidate gene or genome-wide association studies.

Measures of linkage disequilibrium

Briefly, the two most commonly used measures, $|D'|$ and r^2 are introduced (for a more detailed discussion see chapters 3 and 5). If there are two loci, each with two alleles (A and a at the first locus and B and b at the second locus), and the frequencies of alleles A and B are p_A and p_B, respectively, and the frequency of alleles A and B together on the same chromosome is p_{AB}, then the covariance between the loci is $D = p_{AB} - p_A p_B$, where $p_A p_B$ is the expected value of p_{AB} in the absence of allelic association (or coupling), so under linkage equilibrium $D = 0$. When D is positive, p_{AB} has a maximum value equal to the smaller of p_A or p_B and therefore the maximum value of D is the smaller of $p_A(1 - p_B)$ and $p_B(1 - p_A)$; when it is negative, its maximum value is the smaller of $p_A p_B$ and $(1 - p_A)(1 - p_B)$. The sign of D reflects the chance ordering of the alleles at each locus (hence the use of the absolute value of D', $|D'|$), but can be important in the comparison of LD between the same loci genotyped in different populations (e.g. cases and controls) when alleles have been ordered in the same way.

The LD measure r^2 is the squared correlation, where r scales D by the standard deviations of the allele frequencies at two loci, $r^2 = D^2/\{p_A p_B (1 - p_A)(1 - p_B)\}$ (Hill and Robertson, 1968). In contrast, D' scales D by its maximum value given the allele frequencies:

$$D' = D/\min\{p_A(1 - p_B), p_B(1 - p_A)\} \quad \text{if } D > 0$$
$$D' = D/\min\{p_A p_B, (1 - p_A)(1 - p_B)\} \quad \text{if } D < 0$$

Whenever one pair of allele combinations is absent, $|D'| = 1$ and LD is described as ''complete'' because the allelic association is as high as possible given the allele frequency at each locus. For example, if $p_A = 0.6$, $p_B = 0.1$ and $p_{AB} = 0.1$ (hence $p_{aB} = 0$), $|D'| = 1$, but $r^2 = 0.07$. This situation may have an evolutionary explanation: the second SNP (with alleles B and b) may represent a young SNP that originally existed as a monomorphic site with allele b, in which the new polymorphic allele B first occurred on the background of the common allele (A) at the first locus. In contrast, ''perfect LD'' is when only two of the four haplotypes are observed and can only occur when allele frequencies at the two loci are the same, in this case $r^2 = |D'| = 1$. In this case, alleles A and B are likely to have arisen together and to be of similar evolutionary age. $|D'|$ has range 0–1 regardless of allele frequency (although with small sample size, $|D'|$ is often estimated to be 1 when the minor allele frequency is low), whereas the maximum value for r^2 is the smaller of $p_A(1 - p_B)/(1 - p_A)\, p_B$ and its inverse. Studies which describe the observed LD landscape often quote both $|D'|$ and r^2 which allows an at-a-glance judgment of LD together with the difference in allele frequencies of the coupled alleles. For example, high $|D'|$ and high r^2 means a tendency to the presence of only two haplotypes, and a small difference in allele frequency of coupled alleles; high $|D'|$ and low r^2 implies a tendency to the

presence of only three haplotypes, and different allele frequencies of the coupled alleles; low $|D'|$ and low r^2 means a tendency toward random coupling of alleles and presence of all four haplotypes. Given the allele frequency at one locus p_A and its r^2 relationship with locus B, the lower and upper bounds on the frequency of the coupled allele is $p_A - p_A\,(1 - p_A)(1 - r^2)/\,\{1 - p_A(1 - r^2)\}$ and $p_A + p_A\,(1 - p_A)(1 - r^2)/\{p_A\,(1 - r^2) + r^2\}$ (Wray, 2005). From these relationships we can enumerate the high implicit constraints placed on allele frequency given r^2. For $r^2 = 0.8$, the maximum difference in allele frequency is ± 0.06 which occurs when one locus has allele frequency 0.5. For $r^2 = 0.8$ and allele frequency at one locus of 0.1, the maximum difference in allele frequency at the second locus is only ± 0.02 (Figure 4.2).

Estimation of LD requires knowledge of the frequency of haplotypes. Direct sequencing of chromosomes is costly and impractical on a large scale at present, and so haplotype frequencies are not usually directly observed. An individual with genotypes AA and BB at two loci is known to carry two A-B haplotypes. Similarly, an individual with genotypes AA and Bb is known to carry haplotypes A-B and A-b, but an individual who is heterozygous at the two loci, with genotypes Aa and Bb, may carry haplotypes A-B and a-b or A-b and a-B. When a study sample is genotyped, haplotype frequencies are estimated using the EM (estimation-maximization) algorithm, whereby the allocation of double heterozygous individuals to haplotype classes is optimally allocated (via iteration) based on the known haplotype frequency from individuals who are double homozygotes or are homozygous at one locus and heterozygous at the other.

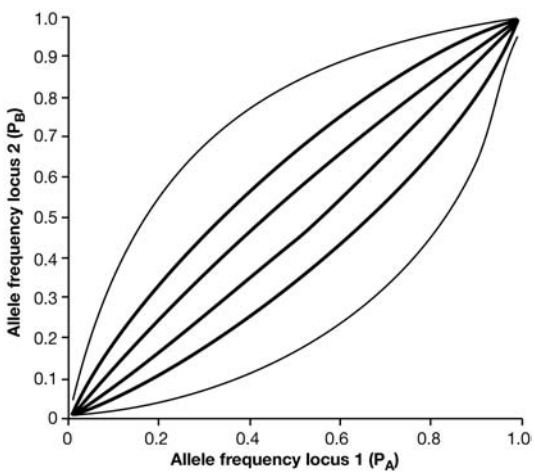

Figure 4.2 Possible range of allele frequencies at two loci given the LD between the two loci (r^2). All possible combinations of allele frequencies are contained within the ellipses for $r^2 \geq 0.8$ (–), 0.5 (–) and 0.2 (–) (from (Wray, 2005)).

Power or sample size in association studies

A key objective in the design of association studies is the determination of the sample size needed in order to be able to detect causal variants of the effect size of interest with a specified power, or if the sample is already collected, determination of the power of the sample to detect a variant of the effect size of interest. Power is the probability (range 0–1) of rejecting the null hypothesis based on the study sample and testing procedure, when in the population the alternative hypothesis is true. One minus power is the false negative rate. In both sample size and power calculations, it is necessary to specify a type I error rate (probability of rejecting the null hypothesis when the null hypothesis is in fact true, the false positive rate), which should be small to account for the level of multiple testing that is anticipated. Often the same sample will be used in association studies of different candidate regions and power calculations should take into account the multiple testing across studies as well as within. Power calculations can be conducted on line e.g. the Genetic Power Calculator (GPC, http://pngu.mgh.harvard.edu/~purcell/gpc/) (Purcell *et al.*, 2003). For a case control study there are a number of input parameters:

i) Frequency of the high risk causal (A allele) variant, which can take values of 0–1.

A reasonable value for this parameter is 0.1 or 0.9 representing either disease risk or disease protective alleles.

ii) Prevalence of the disease in the population. Results are quite robust to any realistic range of prevalence levels.

iii) The genotype relative risk (GRR) of genotype Aa relative to genotype aa and of genotype AA relative to genotype aa, $GRR_{Aa} = P(case|Aa)/P(case|aa)$, $GRR_{AA} = P(case|AA)/P(case|aa)$. Usually in association studies the interest is in variant effects of rather low GRR_{Aa}, in the range of 1.4 to 2. The relationship between GRR_{AA} and GRR_{Aa} depends on the assumptions of the underlying genetic model. For a multiplicative model of A alleles, $GRR_{AA} = GRR_{Aa}^2$ or $(P(case|AA)/P(case|Aa) = P(case|Aa)/P(case|aa))$.

iv) Allele frequency of the genotyped marker and a measure of linkage disequilibrium between the causal variant and the genotyped marker. The GPC requests $|D'|$ as the measure of LD, because any combination of the three 0–1 variables, frequency allele A, frequency allele B and $|D'|$, is possible. In fact, it is more natural to consider the r^2 relationship between the causal and genotyped variants because the sample size of an association study should be increased by a factor of $1/r^2$ to retain the same power if a variant in LD with the causal variant is genotyped rather than the causal variant itself (Pritchard and Przeworski, 2001). As discussed above, the difference in frequencies of alleles at two loci is constrained for a given r^2 relationship between them.

v) The type I error rate, usually 0.05 divided by the number of independent tests or 5×10^{-8} for genome-wide association studies (Risch, 2000).

vi) The number of cases and the ratio of cases to controls (usually 1:1 or >1); the GPC generates the power for this sample size. Increasing the number of controls relative to the number of cases can be a cost-effective way to improve the power of an association study, because controls are usually easier and cheaper to collect (McGinnis *et al.*, 2002).

vii) Desired power. The GPC calculates the required number of cases, given the specified ratio of cases:controls needed to achieve the desired power.

Power of association studies is usually investigated in terms of the GRR of the risk variant (Risch, 2000). For a study sample of 1000 cases and 1000 controls, with desired power 80% and a type I error of 5×10^{-8}, a risk variant with multiplicative allelic action can be detected if it has GRR>1.4 when the genotyped disease risk allele frequency is in the range 0.1–0.9 (Figure 4.3a). The risk variant must have increasingly high GRR to be detected if the risk variant frequency decreases below 0.1 or increases above 0.9 (where the alternative variant is considered to be protective in controls). It is also useful to express the GRR that can be detected in terms of differences in allele frequencies between cases and controls as these are directly observed in case-control studies (Figure 4.3b). These sample sizes are in stark contrast to those required for a linkage study of the same effect size. For example, for a GRR of 1.4 and a multiplicative risk, the required number of affected sibling pairs to achieve the same power in a genome-wide scan is of the order of tens of thousands (Figure 4.1).

Sample collection and population stratification

For any study it is important to consider confounding factors that may result in a spurious association. To this end, control subjects should be ascertained from the same population as the case subjects. For late onset diseases, control samples should have the same age distribution to avoid spurious associations with loci affecting survival to older age. Confounding factors, if known, can be accounted for in the analysis. Indeed, a mixed population of two or more genetic subgroups which differ in their incidence of the disease of

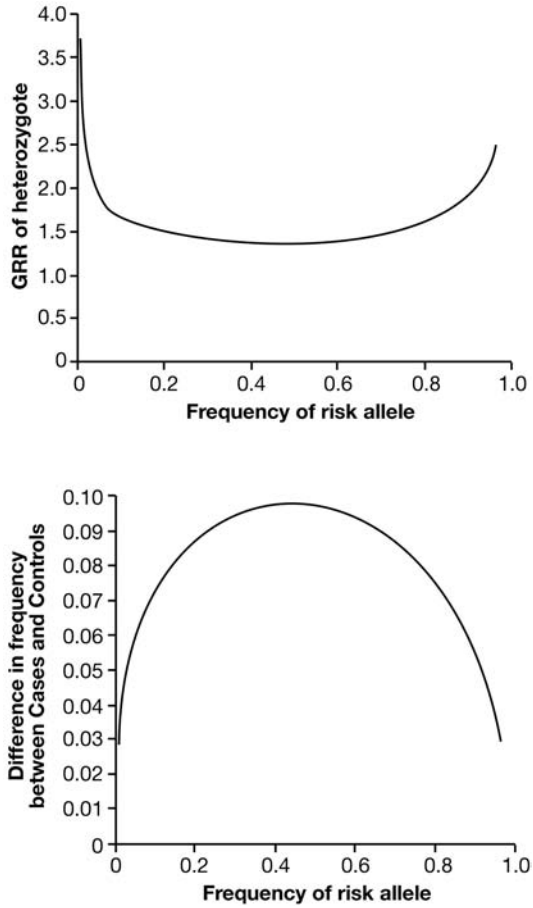

Figure 4.3 Effect size that can be detected in a (genome-wide) association study of 1000 cases and 1000 controls with type I error of 5×10^{-8} and 80% power, assuming multiplicative risk and a disease prevalence of 0.10. (a) Effect size in terms of genotype relative risk of the heterozygote and allele frequency; (b) effect size in terms of the difference in the allele frequency between cases and controls.

interest, may be a powerful population for genetic mapping studies, if genetic markers are available to distinguish the ancestry of chromosomal segments (McKeigue, 2005). However, unknown admixture of genetic subgroups has been a topic of hot debate. Study designs of ascertained probands and their parents have been suggested as a way to guard against problems of unknown admixture. The parent alleles can be partitioned as transmitted to affected offspring (''case alleles''), or not transmitted to affected offspring (''control alleles''). In this way, it is ensured that case and control alleles are drawn from the same population. However, this design to control for admixture comes at a price, as more genotyping is needed to achieve the same number of case and control alleles as in a standard design. Furthermore, it is not always possible, and it is certainly more costly, to obtain family trios. Therefore, case-control studies are still favored, but checks should be made for population stratification by genotyping approximately 20 microsatellite markers or at least 30 SNP markers across the genome (Pritchard and Rosenberg, 1999) and analyzing for stratification, for example using the software STRUCTURE (http://pritch.bsd.uchicago.edu/software.html).

SNP selection

Association studies of a candidate region or other region of interest should aim (unless it is a replication study) to ensure that association is tested using all known variants above a specified frequency either directly or indirectly through the known association between genotyped and ungenotyped variants. After the first reporting of linkage disequilibrium in ''blocks'' across the genome (Gabriel *et al.*, 2002) a plethora of papers emerged in 2003–4 on methods to determine haplotype blocks and to undertake SNP selection for association studies. Many of these methods did not make the connection between SNP selection and subsequent association study analysis. A subtle distinction was made between the notion of SNP tagging and haplotype tagging. In SNP tagging each excluded SNP is tagged by a proxy, which may be a single SNP or multiple-SNP haplotype in high LD with it. In haplotype tagging, SNPs are selected for genotyping so that each haplotype above a given frequency in a haplotype block can be tested for association. The dust has now settled and the

consensus is that SNP tagging is more interpretable when the subsequent association study is analyzed. SNP selection should be based on the squared correlation, r^2 between each unselected SNP and its proxy of genotyped SNP(s). If the genotyped proxy is a single SNP, r^2 is also the measure of LD for the two SNPs. The r^2 measure is used because of its relationship with power as discussed above. This SNP selection method is implemented in freely available software, for example Tagger (de Bakker *et al.*, 2005) which is recommended on the HapMap website (www.hapmap.org). Association analysis conducted subsequent to this method of SNP selection follows naturally, so that association can be interpreted both for each SNP genotyped and for each SNP not genotyped, either through single or multiple SNP haplotype tests. Because of the relationship between r^2 and allele frequencies discussed above, choice of a high threshold for r^2, necessarily means close coupling of allele frequencies between eliminated markers and their single SNP genotyped proxies. Therefore, use of single SNP proxies can result in a high number of SNPs being selected. Reducing the numbers of SNPs selected can be done either by setting a less stringent r^2 threshold or by maintaining the r^2 threshold but considering proxies based on two or three SNP haplotypes. Reducing the r^2 threshold and retaining the property that each unselected SNP is represented by only one other is appealing, because the follow-up analysis requires only single SNP analysis of the genotyped SNPs to be able to make interpretations about all the unselected SNPs. Unfortunately, because of the relationship between r^2 and power, unless the sample size is very large, the power to detect a causal variant may become compromised. It is better for statistical power to retain a high r^2 threshold and use multi-SNP tags than to lower the r^2 threshold (de Bakker *et al.*, 2005). Typically, about 20 to 40% of all available common SNPs are selected for tagging, depending on the exact selection criteria used, representing a substantial cost saving in genotyping.

For any SNP selection procedure there is ample scope to impose other criteria. For example, two SNPs in perfect LD with $r^2 = 1$ are both a perfect proxy for each other. Other criteria can be used to decide which SNP to select for genotyping. For example, positional information (e.g. exonic, promoter), functional information (e.g. non-synonymous base substitutions) or practical information (e.g. ease of assay design) can be included.

The first release of the SNP genotyping of the human genome in 2004 through the HapMap consortium project (HapMapConsortium, 2003; Altshuler *et al.*, 2005) meant that overnight, SNP selection procedures that had been performed bespoke for each association study, taking several months, could be done with a few clicks of the computer mouse. The current HapMap release (number 20) has genotyped nearly 6 million SNPs on 270 people from four separate identified populations (Caucasian, Chinese, Japanese and Yoruban) and genotypes can be downloaded for specified regions or the whole genome. There remains some discussion that SNP selection should be appropriate for the LD of the population under study, in particular the case population. Construction of a study-specific LD map is time consuming, and most likely all studies will use the HapMap, at the very least, as a starting point.

Analysis of association studies

Genotyped markers should be tested for departures from Hardy–Weinberg equilibrium (HWE) in the control population prior to association analysis. HWE describes the expected relationship between allele and genotype frequencies under normal population conditions (see Chapter 3). More often than not, departures from HWE indicate problems with genotyping assays and any such markers must be excluded. True Hardy–Weinberg disequilibrium in the study sample should result in markers in LD with each other all showing departure from HWE. Association analysis of such markers may be interesting but special care is needed for interpretation. Departure from HWE in the case sample

but not in the control sample may reflect association of the marker with disease status.

For a case-control study, association between a disease and alleles of a single SNP can be tested most easily as a 2×2 chi-squared contingency table test with the null hypothesis ''No association between case-control status and allele frequency.'' In this test, the number of each allele type found in cases and controls (obs) is compared to its expected number (exp). For example, the number of A alleles expected in cases $= N_{case}{}^{*}N_A/N_{Tot}$ where N_{case} is twice the total number of cases (the factor of two because each person has two alleles), N_A is the total of A alleles across both cases and controls and N_{Tot} is twice the number of cases and controls in the study. The test statistic is the sum of $(\text{obs} - \text{exp})^2/\text{exp}$ for all four case-control by allele subclasses and, for large sample sizes, follows a chi-square distribution with 1 degree of freedom. Genotype tests can be similarly set up as 3×2 contingency table tests with 2 degrees of freedom, or genotypes collapsed into 2×2 tables for assessment of recessivity or dominance. Association analysis can also be conducted as a logistic regression of the binary y-variable (case = 1, control = 0) on allele or genotype classes. The null hypothesis is ''no association'' and the test statistic is a likelihood ratio test. For the simple scenarios of case-control associated with a single SNP allele or genotype described above, the test statistics obtained via the chi-squared contingency table test and logistic regression are the same. However, the logistic regression framework is general and can be extended to more complex situations, for example to include covariates or to regress on multi-SNP haplotypes. The logistic regression framework can account for uncertainty of estimation of haplotype frequencies via the EM algorithm discussed in the LD section above, the uncertainty in haplotype transmission from parents to offspring in a parent–offspring trio design and uncertainty in transmission resulting from missing parental genotypes. Many free programs are available for analysis of association studies e.g. UNPHASED (Dudbridge, 2003). Association tests of haplotypes from multiple SNPs can be ''global'' in which the null hypothesis is that the distribution of haplotype frequencies is the same in cases and controls, or the test can be for an individual haplotype with null hypothesis that the frequency of a specific haplotype versus all other haplotypes considered together is the same in cases and controls. Testing all individual haplotypes dramatically increases the number of tests performed. It can be reserved for follow-up of a significant global test, but for some underlying genetic models, associated variants may be missed as truly associated individual haplotypes may not generate a significant global association. The size of an associated effect can be described in terms of the increased frequency of associated allele A in cases compared to controls as an odds ratio (OR), $\text{OR} = (N_{\text{case-A}}/N_{\text{control-A}})/(N_{\text{case-a}}/N_{\text{control-a}})$ with 95% CI calculated using the standard error of $\ln(\text{OR}) = \sqrt{\quad}(1/N_{\text{case-A}} + 1/N_{\text{control-A}} + 1/N_{\text{case-a}} + 1/N_{\text{control-a}})$.

OR is an estimate of relative risk and so the terms are used interchangeably in the literature. From the estimate of the genotype frequencies and genotype relative risks (GRR), the population attributable risk (*PAR*) can be calculated to show the proportional reduction in disease incidence if the associated variant was eliminated from the population, $PAR = 1 - \frac{f_o}{P_D}$, where P_D is the incidence of the disease in the population, and f_0 is the probability of genotype aa having the disease,

$$f_0 = \frac{P_D}{(p_A^2 GRR_{AA} + 2p_A p_a GRR_{Aa} + p_a^2)}.$$

Where p_A and p_a are the frequencies of the alleles associated and not associated with the disease. For example a variant of the gene TCF7L2 was recently reported to be associated with type 2 diabetes (Grant *et al.*, 2006). The risk allele has a frequency of 0.26 in the population, giving frequencies of the homozygotes and heterozygotes of 0.07 and 0.38, respectively, assuming Hardy–Weinberg equilibrium. The estimates of the relative risks of the homozygous and heterozygous genotypes were

2.41 and 1.45, respectively. Despite the low relative risks, the fact that the risk allele is common gives a substantial *PAR* of 21%. As hundreds of risk variants with relative risk of this size could exist to account for the total genetic variance observed, it is clear that the *PAR* from different risk loci cannot be added. In fact, *PAR* is a multiplicative measure, so that for the example above the removal of the first locus would reduce the disease incidence to 79% of the original value and the removal of a second locus with the same frequency and relative risks would reduce the disease incidence to 79% of the new incidence, or 62% of the original incidence.

There are four factors which affect the design and interpretation of association studies, namely effect size of the causal variant, allele frequency of the causal variant, allele frequency of the genotyped SNPs, and LD between the causal variant and the genotyped variants. These factors are not independent because, as discussed above, the allele frequencies at two loci determine the range of r^2 that can describe the LD between them (Figure 4.2). The interplay between them has been discussed in terms of $|D'|$ (Zondervan and Cardon, 2004) and r^2 (Wray, 2005).

Multiple testing

The number of tests conducted in an association study can be very large. Frequently the same samples are used in more than one study and the number of tests performed should be summed over studies. Often multiple sets of cases are genotyped as well as multiple markers which may be analyzed as single marker genotypes or alleles or in combination as multiple-SNP haplotypes. Genome-wide association studies of 500k SNPs are likely to become standard in which, if all the SNPs acted independently, one might expect $0.05^{*}500\,000 = 25\,000$ SNPs to show association at the 5% level by chance (type I error). Use of the Bonferroni correction, dividing the significance threshold (α) by the number of tests conducted (n), $\alpha' = \alpha/n$, or the more accurate correction, $\alpha' = 1-(1-\alpha)^{1/n}$ is overly conservative because it is likely that many tests are not independent because of LD. One simple alternative is to estimate the effective number of SNPs from the correlation structure (square root of the r^2 measure of LD) between them and use this to make a Bonferroni correction based on the effective number of tests (Nyholt, 2004).

False discovery rate (FDR) (Storey and Tibshirani, 2003) can be used to set a cut-off for *p*-values so that the proportion of false positives out of all positives (true and false) is controlled. FDR is useful when prior knowledge in experimental design means that many tests are expected to follow the alternative hypothesis (of true association). Therefore, FDR may have a role in replication studies of multiple functional candidate genes, but is less likely to be useful in genome-wide association studies or studies of positional candidate genes in which there is no real prior expectation of association.

Permutation testing can be used to correct for multiple testing and involves random assignment of genotypes to case-control status. This random assignment results in any relationship that may exist between genotype and phenotype in the sample being broken up, so that the resulting permuted sample is from the null distribution of no association. The permuted data is then re-analysed and a comparison is made between the largest test statistic from the permuted sample and the maximum found from the true allocation of case-control status. This is repeated many times (1000 to 500 000 times, depending on the study) and the empirical *p*-value of the most associated observed test statistic is the proportion of permutation test statistics that exceed the observed test statistic. The use of permutation testing for correction of multiple tests should not be confused with the use of permutation testing on a single test basis when the distribution of the test statistic is unknown or deviates from the distribution expected from theory because of relaxation of assumptions. However, the same set of permutations can be used for both purposes (Becker and Knapp, 2004).

It has been argued (Lander and Schork, 1994) that researchers should not just correct for the multiple testing they perform in their own studies but should impose a genome-wide significance level of 5×10^{-8} (Risch, 2000) on every study of a new hypothesis even if only a small subset of the genome is studied because ''the genetics community as a whole is thus conducting a whole-genome scan and the full multiple testing threshold should be applied to any positive result.'' The only real way to get around the problems of multiple testing is to ensure study designs include a replication sample. To this end, association studies addressing new hypotheses are often considered ''hypothesis generating'' and a relaxed significance threshold is used to select variants for genotyping in a replication ''hypothesis testing'' sample in which genome-wide significance levels are not required. This is the rationale behind genome-wide association studies which are usually designed in at least two stages. Rather counterintuitively, joint analysis of the two stages has been shown to be more efficient for detecting associated variants if the first stage included at least 30% of the study samples and if more than 1% of genetic markers are genotyped in the second stage (Skol *et al.*, 2006).

Linkage meets association

A large number of research groups have invested enormous amounts of time and effort in collecting pedigrees (phenotypes and genotypes) for linkage studies. Since these families are a sample from the population, usually enriched for disease prevalence, it seems logical to conduct association analyses on the same individuals (Table 4.1). For example, ascertained affected sibling pairs can become part of a case sample for a case-control design. If sufficient pedigree data are available then combined linkage and association analyses can be performed. The advantage of family-based association studies is that family-based controls can be used to prevent false positives due to population stratification. One popular statistical test for a family-based association analysis is the transmission-disequilibrium-test TDT (Spielman *et al.*, 1993). This test was designed to test for linkage in the presence of association. That is, an allele that shows association with a disease in the population is tested for linkage. The TDT was first suggested for a trio design of sets of families with one affected child and genotypic information on the child and both parents. Subsequently, many other TDT-like tests have been proposed for other family structures, for general pedigrees and for haplotypes. The essence of the TDT is simple and elegant. Consider a SNP with alleles B and b. We wish to test for linkage in the presence of association. A test for association would be to compare the frequency of the B alleles in the children against the frequency in a control sample. However, as shown before, such a test can be spurious if B is associated with a subpopulation in a stratified sample. The trick to get around this is to look at the transmission from heterozygous parents (Bb) to their affected children. If allele B is more frequently transmitted than allele b, then this is evidence for linkage between B and a disease allele, even if the population is stratified. The untransmitted allele is used as a family control, and therefore the test is for within-family association, which is linkage.

Quantitative traits

The focus in this chapter has been on disease phenotypes, which are usually scored in a binary manner (yes or no; affected or unaffected). Gene mapping can, however, also be performed on traits that have more categories, in particular continuously varying quantitative traits. The advantage of working with quantitative traits is that the whole spectrum in the population can be measured. For common disease, many individuals will be at risk but have not yet shown symptoms whereas others may display only mild symptoms. If there was an underlying continuous scale of liability to disease, and if we could measure individuals on that scale,

Table 4.1. Comparison of linkage and association approaches in gene mapping

	Linkage	Association
Sample	Families with known pedigrees Ranging from a small number of large pedigrees with many affected family members to a large number of affected sibling pairs	Disease traits – unrelated cases and controls or family trios allowing construction of pseudo-case (transmitted) and pseudo-control (not transmitted) chromosomes Quantitative trait – population sample, perhaps ascertained to be extremely high- or low-scoring with respect to trait
Sharing of markers between affected individuals	Identity by descent through known pedigree	Identity by state as pedigree is unknown, but with underlying hypothesis that a causal variant arose in a distant common ancestor, or else that the same causal variant arose multiple times
Number of meioses from common ancestor	Small	Large
Length of DNA segment shared between affected individuals with mutation from same common ancestor	Long	Short
Markers of choice	Microsatellites	SNPs
Marker density	Low, spaced <10 cM	High, <1 cM
Marker selection	Based on genetic distance spacing and heterozygosity	Based on LD of the region, SNPs selected to represent the LD of the region and to avoid cost of redundancy of SNPs because of the high LD between them
Effect size that can be detected, which depends on sample size	Generally effect size must be large, odds ratio >3 Quantitative traits: effect size explaining at least 10% of variation	For sample size of 1000 cases and 1000 controls can detect effect size of OR >1.5 with power 80% and significance level 5×10^{-8}, allele frequency of associated variant 0.1–0.9 and disease prevalence 10% Quantitative traits, effect size explaining at least 1% of the variation
Significance needed for acceptance of genome-wide significance	Suggestive: LOD $= 2$ Significant: LOD $= 3$ Replication: LOD $= 1$	Significant: 5×10^{-8} Replication: Bonferroni correction for number of tests which should be small. 0.05/(number of tests)

then linkage and association analyses would be much more powerful. Some quantitative traits may come close by being correlated to liability, for example blood pressure, body mass index and questionnaire based phenotypes in mental health studies.

The principle of linkage and association studies for quantitative traits are much the same as those

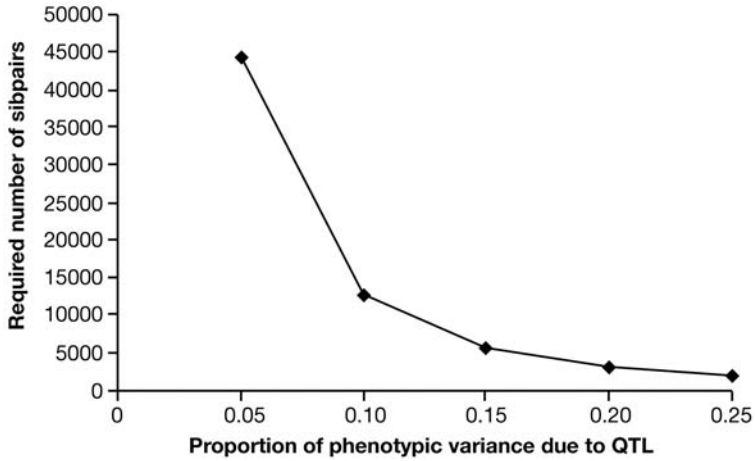

Figure 4.4 Required number of sibling pairs to detect linkage with a quantitative trait in a genome scan, for a power of 80% and a type I error rate of 0.0001 and a sibling correlation of 0.30.

for disease phenotypes, as are the experimental designs. The analysis of data is different, however, because one has to take account of the continuous distribution of the phenotype. For linkage analysis, this is done by estimating the similarity between relatives as a function of the number of alleles they share IBD at a particular locus. If a genomic region harbours a gene or multiple linked genes that cause genetic variation in the population, then relatives that share more of that region will, on average, be more similar in their phenotype. This means that the test for linkage is usually equivalent to testing for significance of a variance component. Unfortunately, the estimation of variance components is imprecise, in particular when compared to estimating mean effects. In the simple case of estimating a phenotypic variance (σ^2) from a sample of n individuals, the standard error (SE) of the estimate is approximately $\sqrt{(2/n)}\sigma^2$. In other words, the SE relative to the population value of the variance is $\sqrt{(2/n)}$. For example, for 1000 observations the SE relative to the true value is $\sqrt{(1/500)}$ or 4.5%. In contrast, the SE of estimating the mean (which is σ^2/n) is much more precise. Most quantitative traits such as blood pressure and body mass index have a coefficient of variation ($CV = \sigma/\mu$) of 10 to 15%. Hence, the SE of estimating the mean as a proportion of the true mean value is $(\sigma/\sqrt{n})/\mu = CV/\sqrt{n}$, and for $CV = 0.15$ and $n = 1000$ this is 0.47%, nearly a 10-fold difference with the estimate of the variance.

For this reason, a random sample of families from the population requires a very large sample size to detect a quantitative trait locus (QTL), just as one would need a very large random sample of families to detect linkage to common disease. In Figure 4.4 the required number of randomly selected sibling pairs is shown to detect a QTL with 80% power at a false positive rate of 1 in 10 000 (which is the appropriate rate for a genome scan). The numbers were calculated from a web-based power calculator (Purcell *et al.*, 2003). Figure 4.4 shows that >10 000 pairs are needed to detect a QTL that explains 10% of the total variance for a trait with a sib correlation of 0.3. If the sibling correlation is entirely due to additive genetic effects, then the total proportion of variance due to genetic factors is 60%. Hence, a QTL that explains 10% of the total variance needs to explain about 17% of all genetic variance. There are few examples of QTLs that explain such a large proportion of the genetic variance. It should be

clear that selecting random sibling pairs from the population is extremely inefficient with regard to linkage analysis of quantitative traits.

To increase power, researchers working on the genetics of quantitative traits try to enrich their sample for a genotype–phenotype relationship by sampling extreme phenotypes from the population, analogous to sampling families enriched for disease. One design to achieve this is called EDAC (for Extremely Discordant and Concordant), and is usually applied to discordant and concordant sibling pairs. In an EDAC design, the most informative individuals for linkage are selected from a large sample of individuals with phenotypes in the population, and only those selected individuals are genotyped. The EDAC design capitalizes on the fact that most of the information on linkage is contained in individuals with extreme phenotypes. For example, extremely discordant sibling pairs (one with a low scoring phenotype and one with a high scoring phenotype) are much more likely to share less than 50% of their alleles IBD at a QTL than two siblings whose phenotypes are around the average. Similarly, extremely concordant pairs are more likely to share more than 50% of their alleles IBD at a QTL. The EDAC design is efficient with respect to genotyping, because only 100s of pairs need genotyping to achieve sufficient power (compared with >10 000 pairs in a random sample from the population). However, it requires the availability of a large sample of families with phenotypic information. The EDAC design has been successfully applied to neuroticism (Fullerton *et al.*, 2003).

Some success stories

It is not the remit of this chapter to provide a comprehensive list of linkage and association success stories, as a number of these will be covered in the chapters of this book dealing with specific diseases. Nonetheless, a few success stories are described which illustrate the progress made in a range of diseases via a variety of study designs.

Linkage studies for common disease do not usually directly lead to underlying causal variants because the effect of any one allele is too small to result in a mapping resolution at the level of individual genes. However, the BRCA1 and BRCA2 susceptibility genes for breast cancer were first mapped using linkage in highly enriched families (Goldgar *et al.*, 1993) and subsequently identified using association and sequencing studies. A number of variants at these genes have a very large relative risk in carriers. The frequency of these variants varies across ethnic groups, and a number of them are at relatively high frequency in Ashkenazi Jews. Since the frequency of these alleles is relatively rare in the overall population, it is doubtful that a genome-wide association analysis would detect them. This example highlights the value of pursuing rare monogenic variants of common diseases as a strategy to increase understanding of complex disorders (Antonarakis and Beckmann, 2006).

In 2001 three papers reported on the association between the gene *NOD2* and Crohn's disease (an inflammatory bowel disease), however, the paths of the research that led to these independent findings were quite different. The first (Hugot *et al.*, 2001) followed the classic research paradigm of genome-wide linkage study (which identified the pericentromeric region of chromosome 16), fine mapping study (where more microsatellite markers are genotyped in the linkage region) and association study. This study identified three causal mutations, a single base-pair insertion (3020Cins) and two missense variants. In contrast, the other studies (Hampe *et al.*, 2001; Ogura *et al.*, 2001) selected the NOD2 gene as a candidate gene based on position (within the reported linkage region), structural homology to plant apoptosis regulatory and disease resistance genes and known function of NOD proteins in recognizing bacterial components. Both studies identified the single base-pair insertion and reported GRR for the heterozygote of 1.5 and 2.6 and for the homozygote of 17.6 and 42.1 respectively. There is no doubt that genetics research has made a major contribution

to the understanding of inflammatory bowel disease and use of genetic markers for diagnosis, prognosis and as a decision tool for choice of drug treatments is predicted for the near future (Vermeire and Rutgeerts, 2005).

Results of genome-wide association studies are only now starting to be published, but a surge of results is expected over the next couple of years. Initial results are promising. For example, the first large scale association studies for schizophrenia (Mah *et al.*, 2006) and Parkinson's disease (Maraganore *et al.*, 2005) both identified variants in genes not previously considered for these diseases but which were interesting functional candidates. A genome-wide association study of age-related macular degeneration (AMD, the leading cause of blindness in the elderly) found (with a very small sample size) association ($p < 10^{-7}$) with a polymorphism in the gene for complement factor H. This gene was simultaneously studied in positional candidate gene association studies (Edwards *et al.*, 2005; Haines *et al.*, 2005) of genes in the previously reported linkage region (1q31–32). The functional variant was identified as tyrosine-histidine change at amino acid 402, with possession of at least one histidine variant increasing risk to AMD 2–7-fold. The histidine variant is common (frequency about 0.35 in controls) and is estimated to account for up to 50% of population attributable risk (Edwards *et al.*, 2005). In the long term the success of linkage and association studies will depend on the, as yet unknown, genetic architecture underlying complex diseases. Only by doing the studies will we begin to understand the number of genes, the number of variants within genes and the effect sizes of those variants. Evidence to date suggests that the Common Variant Common Disease hypothesis is likely to be valid, at least in part, and identification of common variants will open doors to further our understanding of complex disease. However, it is hard to imagine that the experience of Mendelian disorders will not follow through to complex diseases whereby genes influencing disease are found to have many disease risk variants segregating within the population (Stone *et al.*, 2004). The challenge to develop methodology that can identify as yet unknown disease risk genes which are characterized by rare allelic variants alone is still waiting to be met.

Future considerations and prospects

Is there a future for linkage analysis?

For any disease there have been multiple linkage ''replications'', i.e. reports of suggestive or significant linkage in the same genome region as previous findings from independent studies. Although a number of these replications are real, in the sense that the same loci are predisposing to disease in multiple populations, the practice of selective reporting and the usual absence of formal tests of replication suggests that there will be false positives too. The converse is also true: a report of a LOD score of 3 in one study and ''failure to replicate'' in another study does not by itself exclude the existence of the same loci having effects in both populations. If the power of detection is low, then it may have been fortuitous to detect the locus in the first scan whilst finding the expected result in the second scan. From the theory of chi-square distributions, the standard deviation of LOD scores around an expected LOD score (ELOD) is, approximately, $\sqrt{[2(9.21^{*}\text{ELOD-1})]/4.605} = \sqrt{[18.42^{*}\text{ELOD-2}]/4.605}$. So, for expected LOD scores of 1.0, 2.0 and 3.0 at a particular genome location, the standard deviation across replicate studies is about 0.9, 1.3 and 1.6, respectively. This very large variation, relative to the expected value, explains at least partly the difficulties in drawing inference about replication and non-replication. Given the sample size and putative effect of the trait locus, the variation in test statistics among replicate samples can be calculated and used to decide on (non-)replication. A formal statistical test for replication can be performed when analysing two or multiple datasets jointly (Visscher *et al.*, 2005).

The emphasis in human genetics at present is on association studies, because it is in principle more

powerful in detecting small effects and, to some extent, because linkage studies have not delivered an easy route to causative genes and mutations. However, linkage studies should not be dismissed as yet. When combining resources from several research groups, really powerful sample sizes can be obtained to map reliably and accurately disease loci using linkage, without the risk of false positives due to population stratification. An alternative to a joint analysis of multiple genome scans of the same phenotype is to do a meta analysis, in which statistical inference about susceptibility loci is drawn by accumulating, across all studies, the evidence for linkage for a particular genome region. As seen above, the standard deviation of a LOD score is large, so it is entirely possible that none of the multiple genome scans show suggestive or significant linkage at a true susceptibility locus, whilst the evidence from combining all LOD scores across the scans is overwhelming (Segurado *et al.*, 2003).

Impact on health

A two-year US National Institutes of Health consultation culminated in the publication of *A vision for the future of genomics research* (Collins *et al.*, 2003) which set out a series of ''Grand Challenges'' for the translation of genomics research into biology, health and society. The ''Grand Challenges'' to biology included a comprehensive understanding of the genetic architecture of disease, of the organization of genetic networks and protein pathways and of evolutionary variation. These challenges whilst exciting are at least rather well signposted in terms of the groundwork already laid by research to date. The ''Grand Challenges'' for health and society formally articulated the long-term aspirations of both researchers and the general public alike in their desire that the investment made into genomics research will ultimately translate into benefits in the health of our nations, but the details of the way in which this will be achieved are much less clear.

It is well within the realms of possibility to think of a day when cord blood from each new baby will be used for complete genome sequencing. Each individual would be given a prediction of genetic risk to disease with recommendation on choice of lifestyle to help prevent or delay onset of the diseases. Indeed, in the near future, it is easy to see that on-demand genotyping of hundreds of thousands of SNPs could be performed as part of a routine health check-up just as a test for cholesterol level is routine today. These genetic markers could be used to provide predictions of genetic risk to disease or be used to determine best drug treatments suited to the genetic profile. The benefits realized from genomics on health could be significant but the concurrent impact on society as a whole needs to be addressed with policy and legislation changes implemented in advance.

REFERENCES

Altshuler, D., Brooks, L. D., Chakravarti, A. *et al.* (2005). A haplotype map of the human genome. *Nature*, **437**, 1299–320.

Antonarakis, S. E. and Beckmann, J. S. (2006). Mendelian disorders deserve more attention. *Nat Rev Genet*, **7**, 277–82.

Becker, T. and Knapp, M. (2004). A powerful strategy to account for multiple testing in the context of haplotype analysis. *Am J Hum Genet*, **75**, 561–70.

Chakravarti, A. (1999). Population genetics – making sense out of sequence. *Nat Genet*, **21**, 56–60.

Collins, F. S., Green, E. D., Guttmacher, A. E. and Guyer, M. S. (2003). A vision for the future of genomics research. *Nature*, **422**, 835–47.

de Bakker, P. I., Yelensky, R., Pe'er, I. *et al.* (2005). Efficiency and power in genetic association studies. *Nat Genet*, **37**, 1217–23.

Di Rienzo, A. and Hudson, R. R. (2005). An evolutionary framework for common diseases: the ancestral-susceptibility model. *Trends Genet*, **21**, 596–601.

Dudbridge, F. (2003). Pedigree disequilibrium tests for multilocus haplotypes. *Genet Epidemiol*, **25**, 115–21.

Edwards, A. O., Ritter, R., 3rd, Abel, K. J. *et al.* (2005). Complement factor H polymorphism and age-related macular degeneration. *Science*, **308**, 421–4.

Fullerton, J., Cubin, M., Tiwari, H. *et al.* (2003). Linkage analysis of extremely discordant and concordant sibling pairs identifies quantitative-trait loci that influence variation in the human personality trait neuroticism. *Am J Hum Genet*, **72**, 879–90.

Gabriel, S. B., Schaffner, S. F., Nguyen, H. *et al.* (2002). The structure of haplotype blocks in the human genome. *Science*, **296**, 2225–9.

Goldgar, D. E., Cannon-Albright, L. A., Oliphant, A. *et al.* (1993). Chromosome 17q linkage studies of 18 Utah breast cancer kindreds. *Am J Hum Genet*, **52**, 743–8.

Grant, S. F., Thorleifsson, G., Reynisdottir, I. *et al.* (2006). Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet*, **38**, 320–3.

Haines, J. L., Hauser, M. A., Schmidt, S. *et al.* (2005). Complement factor H variant increases the risk of age-related macular degeneration. *Science*, **308**, 419–21.

Hampe, J., Cuthbert, A., Croucher, P. J. *et al.* (2001). Association between insertion mutation in NOD2 gene and Crohn's disease in German and British populations. *Lancet*, **357**, 1925–8.

HapMapConsortium. (2003). The International HapMap Project. *Nature*, **426**, 789–96.

Hartl, D. L. and Clark, A. G. (1997). *Principles of population genetics*. Sunderland: Sinauer Associates.

Hill, W. G. and Robertson, A. (1968). Linkage disequilibrium in finite populations. *Theor Appl Genet*, **38**, 226–31.

Houwen, R. H., Baharloo, S., Blankenship, K. *et al.* (1994). Genome screening by searching for shared segments: mapping a gene for benign recurrent intrahepatic cholestasis. *Nat Genet*, **8**, 380–6.

Hugot, J. P., Chamaillard, M., Zouali, H. *et al.* (2001). Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature*, **411**, 599–603.

Lander, E. and Kruglyak, L. (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet*, **11**, 241–7.

Lander, E. S. and Schork, N. J. (1994). Genetic dissection of complex traits. *Science*, **265**, 2037–48.

Mah, S., Nelson, M. R., Delisi, L. E. *et al.* (2006). Identification of the semaphorin receptor PLXNA2 as a candidate for susceptibility to schizophrenia. *Mol Psychiatry*, **11**, 471–8.

Maraganore, D. M., de Andrade, M., Lesnick, T. G. *et al.* (2005). High-resolution whole-genome association study of Parkinson disease. *Am J Hum Genet*, **77**, 685–93.

McGinnis, R., Shifman, S. and Darvasi, A. (2002). Power and efficiency of the TDT and case-control design for association scans. *Behav Genet*, **32**, 135–44.

McKeigue, P. M. (2005). Prospects for admixture mapping of complex traits. *Am J Hum Genet*, **76**, 1–7.

Morton, N. E. (1955). Sequential tests for the detection of linkage. *Am J Hum Genet*, **7**, 277–318.

Nyholt, D. R. (2004). A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet*, **74**, 765–9.

Ogura, Y., Bonen, D. K., Inohara, N. *et al.* (2001). A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature*, **411**, 603–6.

Pritchard, J. K. and Przeworski, M. (2001). Linkage disequilibrium in humans: models and data. *Am J Hum Genet*, **69**, 1–14.

Pritchard, J. K. and Rosenberg, N. A. (1999). Use of unlinked genetic markers to detect population stratification in association studies. *Am J Hum Genet*, **65**, 220–8.

Purcell, S., Cherny, S. S. and Sham, P. C. (2003). Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics*, **19**, 149–50.

Risch, N. J. (2000). Searching for genetic determinants in the new millennium. *Nature*, **405**, 847–56.

Segurado, R., Detera-Wadleigh, S. D., Levinson, D. F. *et al.* (2003). Genome scan meta-analysis of schizophrenia and bipolar disorder, part III: Bipolar disorder. *Am J Hum Genet*, **73**, 49–62. Epub 2003 Jun 2011.

Sham, P. (1998). *Statistics in human genetics*. London: Arnold.

Skol, A. D., Scott, L. J., Abecasis, G. R. and Boehnke, M. (2006). Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet*, **38**, 209–13.

Spielman, R. S., McGinnis, R. E. and Ewens, W. J. (1993). Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet*, **52**, 506–16.

Stone, E. M., Braun, T. A., Russell, S. R. *et al.* (2004). Missense variations in the fibulin 5 gene and age-related macular degeneration. *N Engl J Med*, **351**, 346–53.

Storey, J. D. and Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proc Natl Acad Sci USA*, **100**, 9440–5.

Vermeire, S. and Rutgeerts, P. (2005). Current status of genetics research in inflammatory bowel disease. *Genes Immun*, **6**, 637–45.

Visscher, P. M., Haley, C. S., Ewald, H. *et al.* (2005). Joint multi-population analysis for genetic linkage of bipolar disorder or ''wellness'' to chromosome 4p. *American Journal of Medical Genetics, Part B Neuropsychiatric Genetics*, **133**, 18–24.

Wray, N. R. (2005). Allele frequencies and the r2 measure of linkage disequilibrium: impact on design and interpretation of association studies. *Twin Res Hum Genet*, **8**, 87–94.

Zondervan, K. T. and Cardon, L. R. (2004). The complex interplay among factors that influence allelic association. *Nat Rev Genet*, **5**, 89–100.

5

Population diversity, genomes and disease

Gianpiero L. Cavalleri and David B. Goldstein

Introduction

The striking success in mapping Mendelian disease genes, coupled with the rapid development of genomic methodologies have generated an initial wave of enthusiasm that progress in understanding the genetics of common diseases might experience similar success. Over the past five years however, this optimism has diminished sharply as the difficulty of the task has become apparent. This difficulty is highlighted by the fact that, to date, very few genes have convincingly been shown to harbor variants that influence predisposition to a common disease (Glazier *et al.*, 2002; Lohmueller *et al.*, 2003).

However, recent progress has been relatively rapid in a number of relevant methodologies for genetic association studies. In particular, there has been considerable progress in making use of linkage disequilibrium (LD) to more efficiently and comprehensively represent human genetic variation in association studies. There now appears reason to be cautiously optimistic that the field may begin to progress more rapidly.

This chapter focuses primarily on factors influencing patterns of human genetic variation, and the implications of the patterns of variation in medical genetics. A central theme is the role of the HapMap project in facilitating genetic association studies. We consider both the use of the HapMap to identify single nucleotide polymorphisms (SNPs) that efficiently represent other SNPs (that is, tagging SNPs), and also the use of the HapMap data to help optimise methods for identifying the causal variants that underlie genotype–phenotype correlations.

Genome-wide LD (linkage disequilibrium) patterning

LD, the non-random association of alleles at sites, is a critical factor in the mapping of genes affecting complex disease (see Chapters 3, 4). LD is affected by both population factors such as demography and selection, and by genetic factors such as recombination and mutation rates. In discussing the forces that shape LD, it is useful to distinguish those that influence patterning across the entire genome (e.g. a population bottleneck), from those that have more localized effects (e.g. selection, local recombination rate).

The complicated demographic history of human populations has played a major role in shaping LD across the genome. Genetic drift (due either to a sharp bottleneck, or due to longer term maintenance of a relatively low population size), admixture (the recent mixing of populations with differing allele frequencies) and inbreeding can all generate LD throughout the genome. Below we discuss in general terms how these are thought to have influenced human LD.

Bottlenecks

It is believed that modern humans emerged from Africa within the last 150 000 years, most likely in

two or more waves, and that colonization of various parts of the world was accompanied by expansions and contractions of population size. Bottleneck events associated with this turbulent history would likely have left signatures on LD patterns across the genome. For example, it is thought that a contraction in European population size during the last glacial maximum (20–30 000 BP) followed by a rapid expansion with the melting of the ice caps, would have resulted in decreased haplotypic diversity and increased LD in the present-day population. Indeed, when 19 randomly selected genomic regions were analyzed, African populations showed markedly less LD than the European populations (Reich *et al.*, 2001). Signatures of bottleneck generated LD have also been shown across coding regions such as the *CD4* locus (Tishkoff *et al.*, 1996) and *PAH* (Kidd *et al.*, 2000).

Admixture

Admixture between two populations with different allele frequencies at two loci will produce an association between these two loci (i.e. LD) in the descendant admixed population (Nei and Li, 1973; Chakraborty and Weiss, 1988). The degree of LD generated between two loci is dependant on the difference in allele frequency at those sites in the parental populations and is shown by Equation (5.1).

$$D_x = m(1-m)\delta_A\delta_B \tag{5.1}$$

Where D_x is LD in the admixed population, m is the proportion of inheritance from one parental population and $\delta_A\delta_B$ are allele frequency differences between the two parental populations

At the time of admixture, LD can be generated between even unlinked sites, but this breaks down rapidly with random mating. On average, LD is broken down by recombination at a rate given by Equation (5.2).

$$D = D_0(1-r)^t \tag{5.2}$$

Where D is the classical measure of LD defined by equation 5.2, D_0 is the ''initial'' LD, r is the recombination rate, and t is the time that has passed since the LD-generating event (i.e. admixture event, in this case).

Thus for unlinked markers after just six or seven generations even a substantial amount of LD introduced by admixture will be reduced to essentially nothing. But for more tightly linked markers, the LD that is generated can persist for considerable lengths of time.

The Lemba are an example of an admixed population. These people, natives of Southern Africa, speak a range of Bantu dialects yet claim Jewish ancestry. This claim has been supported by work showing 68% of Lemba Y chromosome to be of Semitic origin with the remaining 32% of Bantu origin (Thomas *et al.*, 2000). When long-range LD was examined on the X chromosome, it was observed that LD was maintained across genetic distances that were an order of magnitude greater in the admixed Lemba population compared with either of the parental populations (Wilson and Goldstein, 2000).

While the timing of the admixture event in the Lemba is unknown, admixture in African Americans is known to be relatively recent, occurring over the last 20 generations. A simple calculation involving Equation (5.2) would predict that long stretches of admixture-generated LD should still exist across the genomes of the African Americans. Indeed, admixture-generated long-range LD in the African American population has been shown empirically by several studies (Rybicki *et al.*, 2002; Collins-Schramm *et al.*, 2003). Indeed, Collins-Schramm *et al.* (2003) showed significant levels of LD in an African American population to exceed distances of 15 megabases.

Admixture mapping

Given the higher levels of LD in admixed populations, researchers are keen to take advantage of this for gene mapping purposes. But admixture

presents us with a double edged sword. On the one hand, the allelic association between unlinked loci generated by admixture, although short lived (see Equation (5.2)) can lead to spurious association (Pritchard and Rosenberg, 1999). But, in populations in which LD between unlinked loci does not exist, higher levels of LD would allow genetic mapping with a coarser map than would be possible in parental populations. For example, it has been estimated that in the region of 150–300 000 markers would be necessary for whole genome haplotype-based association studies in the White population (Gabriel *et al.*, 2002; Goldstein *et al.*, 2003), this figure could be reduced to as little as a couple of thousand given the range of LD in a recently admixed population such as the African Americans.

If a disease-causing locus is known to differ in frequency across two populations, then there will be an over-representation of ancestry from whichever population has the higher proportion of disease causing alleles. By taking advantage of long-range LD generated in the admixed population, it is possible, by selecting informative SNPs that allow assessment of ancestral origin across the genome, to identify efficiently the region(s) associated with the disease. This method, known as MALD (mapping by admixture linkage disequilibrium), offers a saving over haplotype-based association testing as fewer markers need to be examined. Smith *et al.* (2004) have recently published a panel of just over 3000 SNPs which are sufficient to differentiate African versus European chromosomes across the entire genome. The same group simultaneously published a Bayesian approach to infer ancestry at each point across the genome (Patterson *et al.*, 2004). These two studies have effectively evolved the study of admixture mapping from theory to practice.

It must be noted however, that as MALD mapping can only succeed under strict conditions concerning the genetic control of the trait under study (Smith *et al.*, 2004; Patterson *et al.*, 2004), its application is restricted to a limited number of diseases. Also, the interval defined by admixture based mapping tends to be very large, in the region of several centiMorgans. As such, MALD mapping is very much a coarse method of localizing causal variants for complex disease.

Localized LD patterning

Having discussed factors that influence the shaping of LD across the genome as a whole, we shall now focus our attention on factors that influence LD on a more localized level, the two most important being recombination and selection.

The average effect of recombination is to reduce linkage disequilibrium. This can be readily seen in terms of the dynamic for haplotype frequency change where recombination reduces the frequency of haplotypes that are in excess.

$$X1' = X1 - rD$$
$$X2' = X2 + rD$$
$$X3' = X3 + rD$$
$$X4' = X4 - rD$$

Let alleles at the first locus be *A/a* and at the second *B/b*, then $X1$ is the frequency of the ''coupling'' haplotype *A/B*, $X2$ the frequency of *A/b*, $X3$ is the frequency of *a/B*, and $X4$ is the frequency of *ab*, r is the recombination fraction and $D = X1X4\text{-}X2X3$

Thus when D is positive, $X1$ and $X4$ are in excess, and recombination will reduce the frequency of these ''over-represented'' haplotypes.

Recombination and the shaping of localized LD

Recently, the study of the rate and localization of recombination events has been a subject of intense investigation. Although there was little doubt that the genome harboured regions of intense homologous recombination such as the β-globin gene cluster (Chakravati *et al.*, 1984; Jeffreys *et al.*, 2000), it was generally accepted that, these limited examples apart, recombination was for the most

part a random event. As such, LD patterning was thought to be of little use in mapping complex disease genes. In fact a theoretical study assuming a homologous rate of recombination estimated that useful LD (set at $d^2 > 0.1$) would not extend beyond 3 kb (Kruglyak, 1999).

Following this study however, empirical data accumulated that was very sharply inconsistent with these predictions. First, Reich *et al.* (2001) showed that relatively high levels of LD extend over much longer tracts, with the so-called average half-life (the distance at which the average D′ is less then 0.5) extending approximately 60 kb in Europeans. This observation was later recapitulated in a number of studies, most notably in those directly responsible for the idea of ''blocks'' of linkage disequilibrium (Daly *et al.*, 2001; Johnson *et al.*, 2001).

Before the October 2001 issue of *Nature Genetics* introducing the idea of blocks, most studies tended to focus on the average properties of LD. This reflected an implicit assumption that there is a consistent average decay of LD with distance, although with some noise. Daly argued instead that LD changes more abruptly and proposed non-random recombination events as the mechanism behind the structuring. Such abrupt changes in LD could, it was argued, be stochastic (Anderson and Slatkin, 2004) but the observations of Jeffreys suggested that the structured patterning of LD might often correspond to recombination hotspots (Jeffreys *et al.*, 2000; 2001).

The idea of haplotype blocks also led to the proposal of haplotype tagging SNPs, that is SNPs to represent each of the haplotypes found within a region of limited diversity (Johnson, 2001).

There is now a clear consensus in the community that an assumption of homogeneous recombination rates is untenable, and that empirical studies should consider local changes in LD patterns, and not averages across multiple regions. This marks a critical conceptual change in the field, and can be traced to the 2001 issue of *Nature Genetics*. Nevertheless, a number of questions remain about the real pattern of LD, and its causes.

The strongest argument, that hotspots drive discontinuities is surely the first argument provided by Jeffreys *et al.* (2001), of a clear correspondence between experimentally verified hotspots and LD breakdown (for a review of the hotspot-like distribution of recombination see Kauppi *et al.* (2004). Recently, more complicated methods of inferring recombination rates from LD data have generally supported the existence of hotspots throughout the genome. After analysing genome-wide data, McVean *et al.* (2004) estimated that 50% of all recombinational events occurred in 10% of the sequence.

Selection and the shaping of localized LD

The second major factor shaping localized LD patterns is selection. The recent history of *Homo sapiens*, in particular the post-Neolithic period, has seen modern humans exposed to diverse and changing environments. The development of farming some 8–12 000 BP dramatically changed the diets of many humans, and greatly increased population density and mobility. Various kinds of selection would surely have changed during this time both because of diet and the effect of changes in mobility and population density on infectious disease.

Most forms of selection seen in nature are frequency-independent. Such selection can be either directional or balancing. Directional selection drives a variant either to extinction or fixation whilst balancing selection favours maintenance of heterozygosity. Less common forms of selection such as frequency-dependant selection (selection in which fitness is a function of gene frequency) will also maintain variation in a population in a similar manner to balancing selection. Given the driving force behind selection, the ability to detect its genetic signals would help in understanding the forces that have shaped human traits.

At the time of mutation, a newly created allele is in complete LD with all variants on the haplotype on which it arose. Under neutral selection, a long

period of time is taken for such an allele to reach a high frequency. Generally speaking, the older the age of an allele, the greater the chance of a recombinational event occurring and thus the greater the decay of LD. In the case of positive selection, however, a beneficial allele rapidly reaches high frequency and in turn maintains LD over long distances. Any allele on the same haplotype also increases in frequency as it effectively ''hitchhikes'' with the selected variant through LD.

The signal of selection in population genetic data will depend principally on the strength of selection and on how long ago the selection took place. At the sequence level, selection can be identified by an excess of rare variants (due to a selective sweep followed by accumulation of new rare variants), an excess of common variants (as might be caused by balancing selection) or large allele frequency differences between populations (indicative of differential effects of selection in different populations causing an increase in allele frequency in one but not the other). A battery of statistical tests can be applied to detect these signatures of selection (reviewed in Bamshad and Wooding, 2003) with two examples being Tajima's D test (used to identify selection within populations) and F_{ST} (used to detect selection from patterns of differentiation among populations). Critically, these tests appear only to have power of detection in the case of strong selection.

Whilst it has long been known that selection would generate LD in the vicinity of a selected locus, a lack of empirical data has made it difficult to examine this question in detail. By examining the extent of decay of LD along haplotypes of a gene or region of interest it is possible to estimate the age of the haplotype and based on its frequency, its degree of selection. The presence of a long-range haplotype observed at a high frequency would be indicative of the presence of a mutation that has increased in frequency faster then expected under the null hypothesis of neutral evolution. This ''extended haplotype homozygosity'' (EHH) method, developed by Sabeti *et al.* (2002) was tested on core haplotypes from the *G6PD* and *TNFSF5* genes and identified haplotypes under selection in both genes. Classical tests lacked the power to detect selection across the same data highlighting the advantage of LD-based methods not only to detect, but also to quantify selection across haplotypes (Sabeti *et al.*, 2002).

It is currently unclear how important selection is in shaping human LD, though it does seem that strong cases have been made for a few genes (e.g. *CCR5*, *G6PD*, *TNFSF5*, *LCT*), and there is little reason to think these are the only examples.

In summary, the debate over the causes and dynamics of the patterning of LD across the genome is ongoing and will continue to evolve and develop. However, this debate should not detract from the key point that LD has block-like properties and this structure has important implications for disease mapping by LD.

HapMap and tagging SNPs

Concurrent with the discovery that LD was largely block-like (Reich *et al.*, 2001; Daly *et al.*, 2001) came the concept that a set of informative SNPs could represent, or ''tag'', each of the common haplotypes in regions of low haplotype diversity (Johnson *et al.*, 2001). The two concepts merge because regions of block-like LD typically have low haplotype diversity as little haplotype shuffling has occurred due to recombinational events.

Since its conception in the latter half of 2001, a variety of different methods of tagging SNP (tSNP) selection have been developed. Broadly speaking, they fall into two groups, those based on capturing the maximum haplotypic diversity in a region (diversity based) and those based on capturing the maximum association between the tagging and tagged SNP set (association based). A number of detailed reviews of tagging methods have recently appeared (e.g. Weale *et al.*, 2003; Goldstein *et al.*, 2003), hence our discussion here is only introductory. Overall, we favour tag selection methods that focus on the r^2 value for tSNP selection as it allows quantification of loss of

power (as determined by n/r^2 where n is the number of individuals in which the variant was typed, see Pritchard and Przeworski, 2001).

One example is the haplotype r^2 method for association-based gene mapping. Briefly, this method assesses the proportion of variation of a ''tagged'' SNP set that is explained by the haplotypes generated by the tSNP set in a logistic regression (see Weale *et al.*, 2003; Goldstein *et al.*, 2003; Chapman *et al.*, 2003; Clayton website: http://www.gene.cimr.cam.ac.uk/clayton/software). In this way, haplotypes or combinations of haplotypes defined by the tSNP set, can be used to capture tagged SNP diversity. This is a much more efficient approach than selecting tSNPs based on pair wise r^2 values, in which tSNPs are selected with matching of allele frequencies to tagged SNPs.

Up until now, tagging a gene or region was an expensive and time-consuming task that required extensive re-sequencing and genotyping. The availability of the HapMap dataset suddenly makes tagging a much more attractive approach.

Tagging with HapMap data

The HapMap project was conceived with the aim of producing a SNP genotypic dataset with an initial density of 1 SNP/5kb in 3 ethnic groups; Africans (Yorubans), Europeans (CEPH), and Asians (Han Chinese and Japanese) (Int. HapMap Consortium 2003; Altshuler *et al.*, 2005). The basic idea is that this will provide validated markers, and that the genotypic data will allow identification of tSNPs.

The HapMap website (www.hapmap.org) offers a user-friendly interface allowing selection of any region of the genome of interest in a manner similar to previously established genome browsers (UCSC, NCBI etc). Genotypic data is generated from 30 trios of a specific ethnic origin and can be downloaded either as a batch, or for any defined region of the genome.

The first step in the tag selection routine we employ is to infer haplotypes from the genotypic data. One approach for this is using one of the EM-related haplotype inference packages, such as PLEM (available from: http://www.people.fas.harvard.edu/~junliu/plem/click.html). An implementation using the partition-ligation approach is important in our experience because it greatly helps in the analysis of larger datasets. Once reliable haplotypes have been inferred, we are ready to begin the process of tag selection.

Various packages are available for tag selection. Our package, TagIT, is implemented with the Matlab programming environment (available from http://www.mathworks.com). TagIT allows tSNP selection using a wide range of variations of either diversity or association based methods (all are summarized in the user guide). The program, employing an exhaustive process, searches all combinations of sets of tSNPs of a given size in order to identify the best set, using the conditions defined by a given function. TagIT also offers functions allowing the calculation and display of various measures of LD.

Whilst the concept of haplotype tagging emerged from the observation that blocks of LD exist across the genome, the location or presence of block-like LD should not limit the application of the method. In fact, our observations suggest that limiting tagging to within blocks restricts the effective range of tags as cross-block LD cannot be exploited (Goldstein *et al.*, 2003). We thus suggest that, should a region of interest contain an apparent block-like structure, tag selection should be attempted across the region as a whole, rather then by subdividing and selecting tags within each block. We must caution however, that inference of haplotypes over a large region of high haplotype diversity can introduce substantial error and tSNP selection and testing can have computational issues all of which compromise the efficacy of the process. As such, it is important to assess LD patterns across the candidate region as visualization of patterning of LD can help distinguish obvious places to subdivide a region of interest, should tagging not be feasible across the region as a whole.

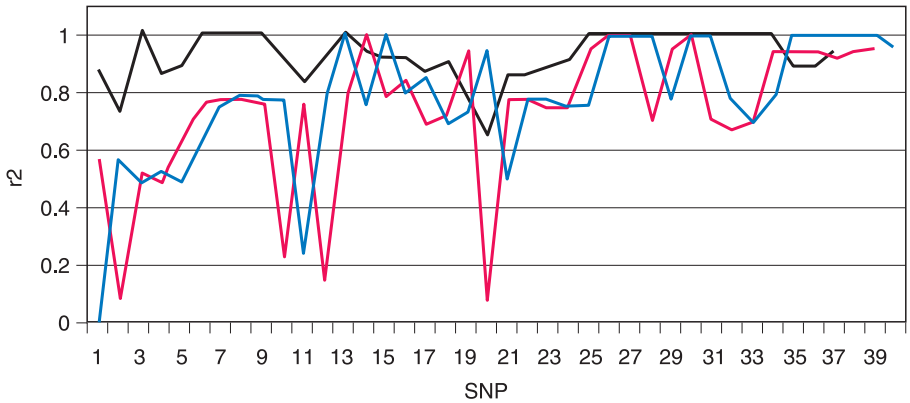

Figure 5.1 Tagging with 3 tSNPs across the *SCN8a* gene.

SNP density and minor allele frequency are central issues in tagging SNP selection. The coefficient of determination from the regression model (referred to here as haplotype r^2) involving the tagging and tagged SNPs provides a formal measure as to how well each individual SNP is being captured by the set of tSNPs. Increasing the number of tags improves performance, however, whilst improvement in performance increases asymptotically, cost does not. A compromise is thus required and a consensus seems to be emerging that the haplotype r^2 (Hr^2) above 0.8 is acceptable (Weale *et al.*, 2003; Goldstein *et al.*, 2003; Carlson *et al.*, 2003), implying that increasing the sample size to $n/0.85$, would be comparable to exhaustive typing.

Example of tSNP design using *SCN8a*

We now provide a detailed illustration of these approaches using one of our studies of the *SCN8A* gene. The first step is to download and format the genotype data from the HapMap project website. An excel spreadsheet provides an efficient platform for data formatting, the principal steps being to order samples as trios, and to code genotypes according to the haplotype inference program used. Once haplotypes have been inferred, the researcher can proceed to tSNP selection.

Tagging a region requires the examination of Hr^2 values for tSNP sets of increasing size. Figures 5.1 and 5.2 illustrate this point using the example of *SCN8a*. An increase in the number of tSNPs from 3 to 6 brings all loci above the threshold Hr^2 of 0.8.

Low frequency variants, although in high LD with surrounding variants, typically tag poorly due to their frequency. But in the case of the *SCN8a* dataset presented here, one variant, SNP12 has a minor allele frequency of 2.5% yet is captured completely by the tSNPs ($Hr^2 = 1$).

The tagging method, when applied to incomplete data (such as that generated by the HapMap project), raises the inherent question of how well unknown variation is being captured. For example, if a functional variant lies on a branch of the genealogy not defined by the tagging SNP set, the power of association would decrease proportionately with the frequency of the untagged haplotype in the unresolved branch of the genealogy. In order to address this crucial point, we developed the SNP dropping procedure (Weale *et al.*, 2003). The approach entails taking the global set of known SNPs and for each SNP *i* dropping it from the analysis in turn. For each reduced set of *N-1* SNPs new tags are selected, and their ability to represent the dropped SNP *i* is assessed (by Hr^2). In this way, a table of Hr^2 values for dropped SNPs is built, providing a statistical estimate of how well

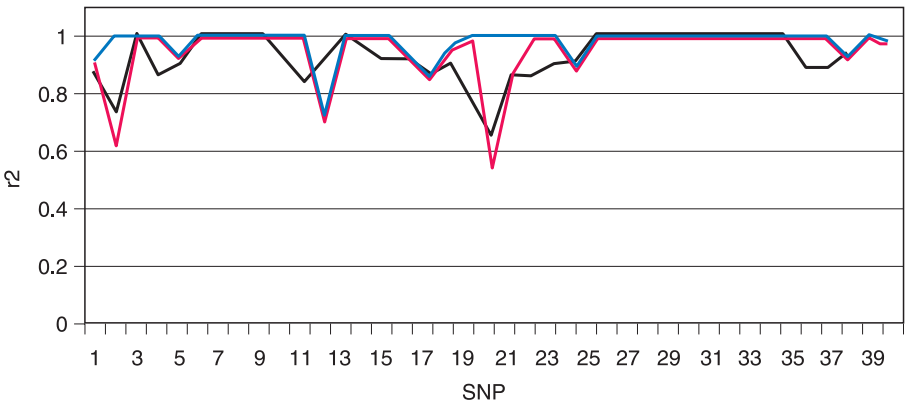

Figure 5.2 Tagging with 6 tSNPs across the *SCN8a* gene.
The solid black lines represent LD values calculated as a sliding window of D′ values (vertical axis) from pairwise comparisons between 3 consecutive SNPs. Blue lines are Hr^2 values. Hr^2 is the coefficient of determination value for that SNP in a linear regression between haplotypes generated by tSNPs and the global SNP set. Red lines are SNP-dropping Hr^2 values. Each value represents the coefficient of determination value for SNP *n* in a linear regression between haplotypes generated by tSNPs (designed on the global SNP set minus SNP *n*) and the global SNP set.

the tSNPs can represent SNPs that are not observed (for example SNPs which are not yet discovered) in the region.

For comparison, we have evaluated tagging sets of 3 and 6 SNPs (Figures 5.1 and 5.2). Given the poor performance of tagging with 3 tSNPs, it can be expected that SNP dropping would, in turn, perform poorly. Increasing to 6 tSNPs improved performance considerably with an average Hr^2 value of 0.95. To evaluate how well undetected variation is likely to be represented, we have found it useful to plot the performance in predicting dropped SNP (e.g. using Hr^2) as a function of the position of the SNP in the gene. With SNP dropping, the dips in Hr^2 values returned indicate poor performance of tags in predicting unknown variation in that region. In the example of *SCN8a* presented here, SNP dropping with 6 tSNPs returns a value of 0.56 at SNP20 (Figure 5.2). Significantly, this dip corresponds to a drop in LD values. Indeed, decay of LD generates haplotype diversity and, as discussed previously, decreases the performance of tagging. For this reason, localized regions of low LD typically correspond with dip in SNP dropping performance.

As discussed previously, there is a trade-off between performance and cost, and hence an acceptable threshold for SNP dropping performance must be defined. At present no consensus exists. Indeed it is difficult to define, as any threshold would depend on the context of the region, however, a value somewhere between 0.5–0.7 as the lower threshold would be acceptable. If, when dealing with incomplete data such as HapMap, SNP dropping indicates a poorly represented region that corresponds with a drop in LD values, similar to the example seen here: the dip in LD values defines the region in which further loci need to be added (either by searching SNP databases or re-sequencing). If however, the region indicated by SNP dropping does not correspond to a drop in LD, then it becomes very difficult to nominate a region in which to focus resequencing, as genealogically informative loci could be located anywhere across the LD block.

In summary, 6 tSNPs are sufficient to satisfy the selection criteria outlined here. These variants capture the vast majority of common variation in the *SCN8a* gene and are ready for application in association studies.

The advent of the HapMap dataset has opened up an array of opportunities in the search for functional common variation across the genome. Up to now, extensive re-sequencing was required in order to characterise the full genealogy of a candidate region, an expensive and time-consuming operation. Whilst considerable work needs to be done to optimize the selection of tSNPs, we believe we are rapidly approaching the point where standardized methods will become available.

Fine localization following initial association

Most attention in terms of the use of the HapMap data has focused on the early steps of selecting and evaluating tagging SNPs. But another potential application relates to one of the major challenges in genetic association, the identification of causal variation responsible for observed genotype–phenotype correlations.

Intuitively, the idea of an associated interval is to define the region around the original associated SNP in which the SNP of putative functionality is likely to be found. The development and optimization of tools to help define these intervals will greatly speed the period between initial association and actual characterization of functional variation. Here we discuss the basic principles, and present a formal method for the definition of an associated interval after detection of the original association.

The idea of an associated interval can be formalized as follows. We consider a variant that has been associated with a phenotype in a ''phenotyped data set''. We denote this variant *m*. Another dataset, ''the LD dataset'', contains information on LD patterns between the associated variant and other linked SNPs. We denote each of these variants *i*. The method involves the application of two separate Bayesian models and compares the association between the associated variant (*m*) and the phenotype to the association between the LD marker variants (*i*) and the associated variant. The first model (Model 0) assumes causality in the genotyped variant and that the genotypic data can explain the phenotypic data via the odds ratio. The second model (Model 1) assumes the genotyped variant is not itself causal but in LD with the causal variant, and that both the odds ratio and the LD value determine its association with the phenotype. The joint posterior distributions of the OR and LD are calculated and the associated interval is assessed via the posterior odds ratio given by: $P(data|Model\ 1)/P(data|Model\ 0)$.

This ratio is calculated for each variant in either direction away from the genotyped variant until a threshold posterior OR is reached at which point the boundary of the associated interval has been formally defined. This method is described in detail in Soranzo *et al.* (2004).

The method can be applied to situations where a variant has been associated with a phenotype of interest in one dataset, but LD information on the region surrounding the association is available from another. Indeed, this method can be applied to any previous association by simply genotyping the associated variant in the same individuals as used in the HapMap project.

Once the associated interval has been formally defined, any variation in the interval, by definition, may be causal. Soranzo *et al.* (2004) took advantage of the fact that tightly associated variants tend to arise on the same branch of the genealogy as the causal variant (Slatkin, 1994). By sequencing across the associated interval in representative chromosomes that are derived from, or ancestral to, the associated variant, it is possible, in a cost-effective manner, to identify candidate causal variants which can then be typed in clinical samples to assess the strength of their association.

We should note that this is different from methods that seek to distinguish polymorphisms, all of which have been typed in the phenotyped data set (see for example plasma ACE concentration and ACE variation; Soubrier *et al.*, 2002). The aim of the associated interval analysis is to determine which polymorphisms need to be typed, and what stretch of sequence needs to be exhaustively re-sequenced. Experience derived from mapping

Mendelian disease genes has placed much emphasis on coding regions in the hunt for functional variation. However, it is becoming more and more apparent that variation in non-coding regions may account for a larger then expected proportion of functional variation. Attention has focused in particular around regions involved in alternative splicing (Pagani and Baralle, 2004). The advent of data provided by the HapMap consortium greatly assists in the definition of such intervals, as SNP selection is not restricted to coding regions and covers the entire genome.

The use of HapMap data to identify evidence for selection

The deluge of publicly available SNP data generated by the HapMap project provides an exciting opportunity to identify regions of the genome under selection. Such work would have the potential to highlight previously unknown regions critical to complex traits thereby opening up new avenues of drug targeting and disease predisposition testing. In this final section we shall briefly discuss the advantages, but also the limitations of the use of HapMap data for this purpose.

The F_{ST} statistic is a measure of how informative a marker is in differentiating populations. Population differentiation can be a signature of localised selection and as such, the F_{ST} statistic can be applied as a tool to detect such signatures. By simply examining the extreme tails of the empirical distribution of a large dataset of genome wide F_{ST} values, one can detect, in principle, candidate regions of localized directional (high F_{ST}), or balancing (low F_{ST}) selection. This method, recently applied by Akey *et al.* (2002) on a dataset of 26 000 SNPs distributed across the genome, was used to identify 174 genes under putative selection. The HapMap dataset, at its initial density of 1 SNP every 5 kb will allow calculation of F_{ST} across three ethnic groups at 600 000 SNPs, an order of magnitude greater then the SNP Consortium dataset used by Akey *et al.*, thus offering the opportunity not only to validate the sites detected by Akey *et al.* but also new ones, given the increased density of the SNP set.

Another tool for the detection of natural selection, also applicable to the HapMap data is the EHH method. To date, this method has only been used on discrete regions of the genome such as *G6PD*, *TNFSF5* and *LCT* (Sabetti *et al.*, 2002; Bersaglieri *et al.*, 2004) and such studies required significant resequencing and genotyping efforts in order to produce datasets of sufficient SNP density. The introduction of HapMap data will not only allow the genome wide application of this method with minimal additional resequencing (to define ''core'' haplotypes) but also allow assessment of significance of results by comparison to empirical data.

However, the use of HapMap data will not always be applicable as the majority of tests of selection assume complete ascertainment of variation. Variation deposited in dbSNP, the database from which HapMap SNPs were selected, is inherently biased towards coding regions.

Conclusions

This chapter has outlined recent progress in our understanding of the patterning of LD across the genome and how this knowledge has catalysed the development of powerful new tools for localising and fine mapping functional variation. The application of methodologies used with such success in mapping Mendelian disease genes has shown that the task of elucidating the genetics of complex disease is extremely difficult and a long road lies ahead. Emphasis must be put on the continued development of technologies such as those described here. Progress in tSNP design combined with the availability of HapMap data now offers the opportunity to undertake powerful association studies in an efficient and well-designed manner. But these are just some examples of how this dataset can be used. The challenge is now for the field of human genetics to develop novel uses of this resource to aid in this undertaking.

REFERENCES

Akey, J. M., Zhang, G., Zhang, K., Jin, L. and Shriver, M. D. (2002). Interrogating a high-density SNP map for signatures of natural selection. *Genome Res*, **12**, 1805–14.

Altshuler, D., Brooks, L. D., Chakravarti, A. *et al.* (2005). International HapMap Consortium. A haplotype map of the human genome. *Nature*, **437**, 1299–320.

Anderson, E. C. and Slatkin, M. (2004). Population-genetic basis of haplotype blocks in the 5q31 region. *Am J Hum Genet*, **74**, 40–9.

Bamshad, M. and Wooding, S. P. (2003). Signatures of natural selection in the human genome. *Nat Rev Genet*, **4**, 99–111.

Bersaglieri, T., Sabeti, P. C., Patterson, N. *et al.* (2004). Genetic signatures of strong recent positive selection at the lactase gene. *Am J Hum Genet*, **74**, 1111–20.

Carlson, C. S., Eberle, M. A., Rieder, M. J. *et al.* (2003). Additional SNPs and linkage-disequilibrium analyses are necessary for whole-genome association studies in humans. *Nat Genet*, **33**, 518–21.

Chakraborty, R. and Weiss, K. M. (1988). Admixture as a tool for finding linked genes and detecting that difference from allelic association between loci. *Proc Natl Acad Sci USA*, **85**, 9119–23.

Chakravarti, A., Buetow, K. H., Antonarakis, S. E. *et al.* (1984). Nonuniform recombination within the human beta-globin gene cluster. *Am J Hum Genet*, **36**, 1239–58.

Chapman, J. M., Cooper, J. D., Todd, J. A. and Clayton, D. G. (2003). Detecting disease associations due to linkage disequilibrium, a class of tests and the determinants of statistical power. *Hum Hered*, **56**, 18–31.

Collins-Schramm, H. E., Chima, B., Operario, D. J., Criswell, L. A. and Seldin, M. F. (2003). Markers informative for ancestry demonstrate consistent megabase-length linkage disequilibrium in the African American population. *Hum Genet*, **113**, 211–19.

Daly, M. J., Rioux, J. D., Schaffner, S. F., Hudson, T. J. and Lander, E. S. (2001). High-resolution haplotype structure in the human genome. *Nat Genet*, **29**, 229–32.

Gabriel, S. B., Schaffner, S. F., Nguyen, H. *et al.* (2002). The structure of haplotype blocks in the human genome. *Science*, **296**, 2225–9.

Glazier, A. M., Nadeau, J. H. and Aitman, T. J. (2002). Finding genes that underlie complex traits. *Science*, **298**, 2345–9.

Goldstein, D. B., Ahmadi, K. R., Weale, M. E. and Wood, N. W. (2003). Genome scans and candidate gene approaches in the study of common diseases and variable drug responses. *Trends Genet*, **19**, 615–22.

Jeffreys, A. J., Ritchie, A. and Neumann, R. (2000). High resolution analysis of haplotype diversity and meiotic crossover in the human TAP2 recombination hotspot. *Hum Mol Genet*, **9**, 725–33.

Jeffreys, A. J., Kauppi, L. and Neumann, R. (2001). Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. *Nat Genet*, **29**, 217–22.

Johnson, G. C., Esposito, L., Barratt, B. J. *et al.* (2001). Haplotype tagging for the identification of common disease genes. *Nat Genet*, **29**, 233–7.

Kauppi, L., Jeffreys, A. J. and Keeney, S. (2004). Where the crossovers are: recombination distributions in mammals. *Nat Rev Genet*, **5**, 413–24.

Kidd, J. R., Pakstis, A. J., Zhao, H. *et al.* (2000). Haplotypes and linkage disequilibrium at the phenylalanine hydroxylase locus, PAH, in a global representation of populations. *Am J Hum Genet*, **66**, 1882–99.

Kruglyak, L. (1999). Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat Genet*, **22**, 139–44.

Lohmueller, K. E., Pearce, C. L., Pike, M., Lander, E. S. and Hirschhorn, J. N. (2003). Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet*, **33**, 177–82.

McVean, G. A., Myers, S., Hunt, S. *et al.* (2004). The fine-scale structure of recombination rate variation in the human genome. *Science*, **304**, 581–4.

Nei, M. and Li, W. H. (1973). Linkage disequilibrium in subdivided populations. *Genetics*, **75**, 213–19.

Pagani, F. and Baralle, F. E. (2004). Genomic variants in exons and introns, identifying the splicing spoilers. *Nat Rev Genet*, **5**, 389–96.

Patterson, N., Hattangadi, N., Lane, B. *et al.* (2004). Methods for high-density admixture mapping of disease genes. *Am J Hum Genet*, **74**, 979–1000.

Pritchard, J. K. and Przeworski, M. (2001). Linkage disequilibrium in humans: models and data. *Am J Hum Genet*, **69**, 1–14.

Pritchard, J. K. and Rosenberg, N. A. (1999). Use of unlinked genetic markers to detect population stratification in association studies. *Am J Hum Genet*, **65**, 220–8.

Reich, D.E., Cargill, M., Bolk, S. *et al.* (2001). Linkage disequilibrium in the human genome. *Nature*, **411**, 199–204.

Rybicki, B.A., Iyengar, S.K., Harris, T. *et al.* (2002). The distribution of long range admixture linkage disequilibrium in an African-American population. *Hum Hered*, **53**, 187–96.

Sabeti, P.C., Reich, D.E., Higgins, J.M. *et al.* (2002). Detecting recent positive selection in the human genome from haplotype structure. *Nature*, **419**, 832–7.

Slatkin, M. (1994). Linkage disequilibrium in growing and stable populations. *Genetics*, **137**, 331–6.

Smith, M.W., Patterson, N., Lautenberger, J.A. *et al.* (2004). A high-density admixture map for disease gene discovery in African Americans. *Am J Hum Genet*, **74**, 1001–13.

Soranzo, N., Cavalleri, G.L., Weale, M.E. *et al.* (2004). Identifying candidate causal variants responsible for altered activity of the ABCB1 multidrug resistance gene. *Genome Res*, **14**, 1333–44.

Soubrier, F., Martin, S., Alonso, A. *et al.* (2002). High-resolution genetic mapping of the ACE-linked QTL influencing circulating ACE activity. *Eur J Hum Genet*, 10, 553–61.

The International HapMap Consortium. (2003). The International HapMap Project. *Nature*, **426**, 789–96.

Thomas, M.G., Parfitt, T., Weiss, D.A. *et al.* (2000). Y chromosomes traveling south: the cohen modal haplotype and the origins of the Lemba – the ''Black Jews of Southern Africa''. *Am J Hum Genet*, **66**, 674–86.

Tishkoff, S.A., Dietzsch, E., Speed, W. *et al.* (1996). Global patterns of linkage disequilibrium at the CD4 locus and modern human origins. *Science*, **271**, 1380–7.

Weale, M.E., Depondt, C., Macdonald, S.J. *et al.* (2003). Selection and evaluation of tagging SNPs in the neuronal-sodium-channel gene SCN1A: implications for linkage-disequilibrium gene mapping. *Am J Hum Genet*, **73**, 551–65.

Wilson, J.F. and Goldstein, D.B. (2000). Consistent long-range linkage disequilibrium generated by admixture in a Bantu-Semitic hybrid population. *Am J Hum Genet*, **67**, 926–35.

6

Study design in mapping complex disease traits

Harry Campbell and Igor Rudan

Introduction

The current rate of advances in genetic technology and statistical methods makes it difficult to discuss study design in mapping complex disease traits in a way that will have value beyond a relatively short time horizon. This chapter considers how knowledge about the nature of complex diseases and traits can inform study design and confines itself to genomic (rather than proteomic or metabonomic) approaches.

Genetic influences on complex traits can be considered in terms of susceptibility to disease (clinical and pre-clinical), susceptibility to differences in natural history of disease (severity, complications and prognosis), susceptibility to different therapeutic responses (efficacy and adverse effects) or in terms of the genetic determinants of normal phenotypic variation in health.

The choices between approaches depend not only on the context of the study, but also on the relative costs of ascertaining families, measuring phenotypes and genotyping. The costs of genotyping have been falling rapidly over the last decade and the trend is for genotyping to be done in a few automated high-throughput centres to maximize efficiency. In contrast, more stringent ethical and data protection legislation requirements have tended to increase unit recruitment costs, since ascertainment and recruitment procedures become more demanding and remain very labor intensive. It is likely therefore that the requirements for very large sample sizes and for large collaborative studies will increasingly involve research groups from countries of intermediate development which can assure high fidelity phenotyping, but at much lower cost than is possible in most industrialized nations. There is also a growing trend towards multidisciplinary collaboration so that genetic data can be complemented by functional studies and set in the context of known molecular pathways.

Research priority setting and investment should consider the burden of disease in developing as well as developed countries and should give emphasis to areas of genetic research in which there is currently a limited ability of public health systems to modify population exposure to non-genetic risk factors. It is also important that there is a critical consideration of where investment in genetic research is or is not appropriate. Thus, for example, in multiple sclerosis, where etiology is largely unknown and treatment unsatisfactory, the greatly increased sibling recurrence risk of disease suggests that genetic research should be supported. In contrast, sib–sib correlations for nicotine dependence are far lower than sib–peer correlations, suggesting that it is principally peer-influenced, indicating that research should focus on interrupting the social transmission of smoking. (Merikangas and Risch, 2003.)

Goals of complex disease trait mapping

The main aims of these studies are:-

1 to understand physiological and pathological mechanisms and pathways leading to disease through the identification of genetic factors underlying these processes. It is hoped that the resulting new knowledge will lead to new treatments and disease prevention strategies.
2 to identify increased genetic susceptibility within individuals where this risk has been shown to be reversible and amenable to reduction. This includes genetic redefinition of disease subtypes, with the goal of explaining differential response to treatments and predicting drug responses or side-effects.
3 to better elucidate the role of environmental risk factors. It is now possible to measure genetic factors with considerably greater precision and validity and at lower cost than for most environmental exposures. Therefore, the identification of novel genetic variants and hence pathways that influence disease risk may suggest which environmental exposures are important in determining disease risk. An important element of this is the use of genetic data in an experimental framework (equivalent to a randomized comparison) to extend observational data by means of ''Mendelian randomisation'' design (Keavney, 2002; Davey Smith and Ebrahim, 2003). This approach controls for most sources of confounding influences on genetic associations and so can be utilised to identify small to moderate environmental effects and to test hypotheses about causal pathways amenable to intervention. The approach does not control for population stratification but this can be minimized through study design or within the analysis (Pritchard *et al.*, 2000; Devlin *et al.*, 2004).

There has been general over-optimism regarding the immediate benefits of improved knowledge of the genetic basis of complex diseases. Public health benefits will not be realized until gene function, and related biochemical and physiological pathways, are understood and molecular targets for new drug discovery identified. The growing realization that new genomic knowledge is the start and not the end of the process is now leading to more emphasis on integrated ''systems biology'' approaches.

Since the greatest public health benefit is expected to follow from the first aim, it will be important to identify and focus on key pathways. Single genes exerting large phenotypic effects, such as cell cycle and mismatch repair genes in colorectal cancer, are likely to be closest to the key, rate-limiting steps in the disease process. From this follows the need to give some priority to identifying moderate-to-large rather than average genetic effects, since these are likely to be concentrated in important pathophysiological pathways. In addition, the complexity of common late-onset disorders suggests that identifying genes with the largest effects, which contribute most to the extremes of the disease or trait distribution, may be the most robust approach, similar in principle to the methods successfully applied to monogenic disorders. For example, the identification of genetic variants accounting for rare monogenic forms of common disease (e.g. breast and colon cancer, Alzheimer disease) has made a substantial contribution towards elucidating disease mechanisms and holds out the potential to lead to therapeutic progress (see Chapter 10). Similarly, knowledge of the genetic basis of a rare Mendelian disorder in which there is a substantially increased risk of a complex disease has provided clues to the etiology of the complex disease. It is less clear whether the identification of variants with small effects on disease risk will have a similar impact.

A counter-argument has been that this approach is important to the families with these generally rare conditions but is of little public health significance. Population attributable fractions (PAF) or risks (which estimate the overall contribution of a risk factor to a disease in a specific population) are often quoted for common variants with small effects in support of their ''public health importance.'' However, unlike prevalent and modifiable environmental exposures

with adverse health effects, PAF values for genetic variants have less clear meaning or public health utility. This is particularly true in complex multifactorial disease in which PAFs typically sum to much greater than 100%, due to interactions among environmental and/or genetic factors. PAFs for genetic variants may retain meaning in terms of the potential for the reduction in drug side-effects, if genetic testing could identify all those genetically predisposed to a certain adverse effect and prescriptions avoided in those found to carry that variant.

Important preliminary steps

This review argues that research investment should target research questions that are tractable with current methods and resources or which are likely to be so with appropriate investment in the foreseeable future. This requires a sober consideration of what is known about the genetic architecture of complex disease so that study designs have adequate power and a realistic chance of success. It is important not to be driven solely by technological advances or, for example, by the commercial drive to develop genetic profiling for complex diseases.

Review of evidence for a genetic effect

Given the large commitment of time and resources required to investigate the genetic basis of complex disease, it is important first to show a clear influence of genetic factors on variation in a complex trait or on disease frequency. In complex disease, a family history is an unreliable guide, as it is often not present, either because the genetic influence is weak or because a polygenic etiology results in disease risk declining exponentially with decreasing relationship to an affected family member. Conversely, if a family history is present, it may result from environmental or ubiquitous factors such as age-associated decline. A first step in establishing a genetic basis is often to demonstrate a correlation between the familial recurrence risk of disease and the degree of relationship, or to measure the heritability of quantitative traits related to complex disease. Thus data from twin, family or adoption studies measuring sib:population risk ratios (λs), heritability data from pedigrees or sets of nuclear families, can all be used to make an estimate of the size of the genetic contribution to disease susceptibility.

Low values for λs or heritability suggest either that investment in genetic research should not be accorded high priority (as in the example above) or that strategies to increase heritability are essential. Strategies for ascertaining study participants so as to improve the ''detectance'' or informativeness of the phenotype for the underlying genotype (Weiss and Terwilliger, 2000) are critically important in common diseases (Table 6.1). Detectance is defined as the probability of carrying a specific disease or trait risk genotype given the particular clinical or subclinical phenotype, which is often low in common multifactorial diseases (Figure 6.1). It can be increased by selecting a subgroup of disease with higher heritability for initial studies.

High values of λs or heritability provide summary evidence for a large total genetic effect but can be consistent with low individual locus or variant specific heritabilities and thus low detectance and study power if many loci or variants contribute to the genetic effect (as may be the case with adult height) (Figure 6.1). If there has been a lack of success in adequately powered linkage studies despite high λs values, it suggests that each individual variant accounts for only a low proportion of the total genetic effect.

In practice, heritability estimates often do not reliably identify which phenotypes are influenced by genes with major effects (Weiss and Terwilliger, 2000) and therefore need to be used cautiously in making decisions. In addition, results are specific to a given population with a particular pattern of environmental exposures. Heritabilities may vary with age, showing a steep decline as age of disease-onset rises, implying a generalised aging process that is not influenced by genetic variation. Even if the total genetic variance is constant or actually

Table 6.1. Strategies to improve detectance, or informativeness of the phenotype for the underlying genotype

- *Familial correlation* (e.g. family history of disease or, for complex traits, sib pairs who are correlated for risk factors such as elevated cholesterol or blood pressure).
- *Severe disease* (e.g. early-onset adult cancers or heart disease).
- *Endophenotypes* (measures can be chosen that minimize the effect of environmental variation e.g. cholesterol ratios (total/HDL, LDL/HDL) are valid risk factors in populations with widely differing mean cholesterol levels).
- *Genetic subgroup* (e.g. colorectal cancer subjects without mismatch repair defects or HLA-matched type 1 diabetics).
- *Low environmental risk* (e.g. chronic obstructive pulmonary disease or bronchial cancer in non-smokers).
- *High environmental risk* (e.g. smokers who develop lung cancer. This assumes that environmental exposure plus a high risk genotype results in interaction and increased susceptibility. As argued in the text, the signal:noise ratio here is very low so that it is generally more efficient to identify susceptibility factors in low exposure groups and then test for interactions).
- *High prevalence group* (e.g. type 2 diabetes mellitus in Micronesions or Mexican Americans (rural versus urban), hypertension in black Africans, primary open-angle glaucoma in blacks and Caribbeans).
- *Less affected sex* (e.g. in systemic lupus erythematosus there is a 9:1 female:male sex difference in prevalence, and increased recurrence risks in relatives of male compared with female subjects, implying stronger genetic effects).
- *Population isolate* (e.g. North American religious sects, isolated populations with ''small founder pools'', or young populations in which genetic associations are substantially stronger (larger regions of linkage disequilibrium) than in large urban populations).
- *Early age-of-onset* relative to the population norm. (e.g. breast or colorectal cancer, heart disease. Age-of-onset may be a more sensitive indicator of a significant genetic component than family history and a familial correlation in age-of-onset even more useful when recurrence risks are low. Early-onset disease tends to be associated with higher heritability, due to reduced environmental and stochastic (e.g. aging) components of variance found at older ages).

Source: Wright *et al.* 2002 (with permission of authors)

increases with age, heritability estimates can decline due to rising non-genetic sources of variance, and so can be misleading. Finally, heritability may vary depending on whether or not dominance effects are considered.

Realistic estimation of study power

Power is governed by the proportion of phenotypic variance in the population that is explained by the genetic variant. Since this is typically low in complex disease, there is a need to achieve large sample sizes in almost all study designs. There is thus an increasing need to channel research funding and effort to large national or international collaborations between government, pharmaceutical industry, medical research charities and universities to achieve sufficient sample size. Lack of power due to unrealistic assumptions about the underlying genetic architecture is the most likely reason for disappointing results from genetic studies of complex disease, in both affected sib pair (Freimer and Sabatti, 2004) and association studies (Zondervan and Cardon, 2004). Collaborations that can combine data for joint analysis and meta-analysis, coupled with the use of higher density marker scans, may be more successful in the future. Although there has been much discussion of the dangers of false positive results, due to a combination of multiple testing without appropriate adjustment of significance levels, coupled to publication bias (Cardon and Bell, 2001), the problem of false negatives represents a more serious challenge to future gene discovery efforts. Research funding agencies and journals could facilitate such collaborations by encouraging joint funding or publication agreements.

An important attribute of complex diseases is the unpredictable relationship between genotype and

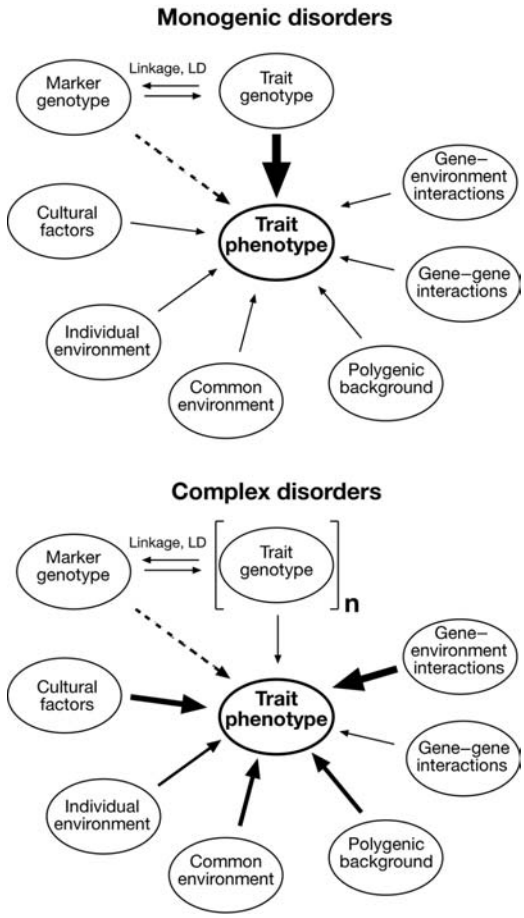

Figure 6.1 Schematic model showing the relationship between phenotype, genotype and ascertainment in complex genetic disorders (Weiss and Terwilliger, 2000). In simple monogenic disorders, the trait phenotype is strongly influenced by a specific trait genotype, although is secondarily influenced by environmental polygenic background and other influences. In this case, the probability of the trait susceptibility genotype given such a phenotype (detectance) is high. In complex disorders, the trait phenotype is strongly influenced by many different genetic, environmental and cultural factors (with interactions indicated by arrows) so detectance is low. Genetic factors may include many loci of small or large effect, and polygenic background. Marker genotypes are near to (and hopefully correlated with) one or more susceptibility genotypes that influence the phenotype. Genetic epidemiology tries to correlate marker genotypes with the trait phenotype in order to localize susceptibility (trait) genotypes. Detectance is a function of the ascertainment scheme. For example, in contrast to the situation with a rare monogenic disorder where detectance is high, additional ascertainment schemes are required in common multifactorial disorders, since the phenotype is influenced by many genetic and environmental factors (Table 6.1). The figure is adapted from Weiss and Terwilliger (2000).

phenotype. Even where a high penetrance genotype predicts the phenotype well, the opposite does not necessarily hold true. Thus, retinitis pigmentosa (RP) genotypes may predict the RP phenotype but the opposite may not hold and detectance is low even in the presence of highly deterministic genetic variants, since there are many different genetic causes of RP (Chapter 32). It is important therefore to improve the detectance of the genotype by strengthening the correlation between phenotype and genotype. Since it has a major impact on study power. This can be achieved by:-

- improving the accuracy and reliability of phenotype measurement
- reducing phenotypic complexity by studying specific subphenotypes, to reduce the number of loci influencing the phenotype and increase the size of the locus-specific effect on the phenotype
- studying endophenotypes (see below) that are more directly related to gene action than categorical disease states
- employing ascertainment strategies to increase the chances of identifying variants of large effect (Wright and Hastie, 2001 and Table 6.1)
- minimizing other effects on the phenotype, for example, by studying populations showing reduced environmental variance (Wright *et al.*, 1999; Peltonen *et al.*, 2000)

- increasing the correlation between genetic markers and trait susceptibility variants, such as by using a higher density of markers.

Genetic architecture of common complex disease: implications for study design

Strategies for identifying genetic variants influencing disease susceptibility are strongly dependent on the relative contributions of large and small genetic effect sizes, of rare and common variants, and on whether these occur at a limited (oligogenic) or a large (polygenic) number of loci. Since the genetic architecture of complex diseases is largely unknown, it is important that analytic methods are robust and perform well under a variety of models and with realistic rather than over-optimistic assumptions.

For example, one strategy is to accept that the genetic architecture of complex disease and the population genetic characteristics of the population under study are unknown, so that the best chance of success is to study a large number of endophenotypes in an intensively phenotyped cohort, to increase the chance of finding favorable study parameters and thus success with at least one quantitative trait (QT). Endophenotypes are subclinical QTs that provide more specific disease-related phenotypes than disease per se and which account for a significant proportion of the genetic component of disease risk. The disadvantage of this approach is that the only means of detecting low frequency variants influencing disease-related endophenotypes is to have very large sample sizes or to sample individuals with extreme trait values. If susceptibility variants are common, such enrichment is unnecessary, but if they are rare, they may be enriched within a disease population, while remaining undetectable even in large general population samples. Regardless of the study population, there is now broad agreement that the study of endophenotypes is desirable. However, the successful expansion of this strategy will require identification of appropriate endophenotypes for diseases in which these are not yet available (such as major psychiatric and neurological disorders and many cancers).

Genetic effect sizes in complex disease: implications for study strategy

The susceptibility genes that are currently being identified in complex diseases are commonly those with effects on disease at the extremes of the phenotypic spectrum, almost all of which are large enough to be detectable by linkage and which show risk allele penetrances in the range 1–10%. The majority of identified QTL individually explain less than 5% of the phenotypic variance and only a small proportion show ''major'' QTL effects accounting for >10% of the variance. For example, currently known genetic variants associated with familial breast cancer account for <2% of the total population risk for this disorder (Pharoah *et al.*, 2002; Chapter 15, this volume). Furthermore, apparently large QTL effects are often explained by the composite effect of numerous clustered variants with small individual effects (Barton and Keightley, 2002).

Genetic effect size (genotype relative risk) may or may not decline with increasing age but its contribution relative to the total phenotypic variance frequently falls as the role of environmental exposures become more important (MacGregor *et al.*, 2003) as has been demonstrated in a number of conditions (Marenberg *et al.*, 1994; Lyons *et al.*, 1998; Wright *et al.*, 2002). This is reflected in the decreasing heritabilities with increasing age-of-onset in breast cancer (Claus *et al.*, 1990) and numerous late-onset traits, as somatic aging and stochastic factors play an increasing role and cause the environmental variance to rise (Wright *et al.*, 2002).

Identifying moderate to large genetic effects

Clues can sometimes be gained from reports of cytogenetic abnormalities (such as chromosomal rearrangements or translocation breakpoints)

in patients or families with several members who have a specific disease (Millar *et al.*, 2005). Alternatively, the search for moderate effect variants in genes known to cause Mendelian diseases with phenotypic overlap with common diseases may be fruitful. For example, resequencing of genes in which severe loss-of-function variants cause Mendelian disorders of lipid metabolism has revealed less severe variants associated with raised levels of high density lipoprotein in a population-based study (Cohen *et al.*, 2004).

Family-based linkage approaches (Chapter 4) are robust to problems of rare variants and allelic heterogeneity, and have proved successful in identifying genes showing large and moderate effect variants. Most linkage studies that have been unsuccessful probably result from inadequate sample sizes and unrealistic assumptions about genetic architecture. This includes linkage peaks that have not been replicated in subsequent studies, due to small to moderate effect sizes (Botstein and Risch, 2003). Familial relative risk estimates (such as λ_S) might need to be partitioned among many genetic loci, so that the power to detect an individual locus by linkage approaches is reduced accordingly. Since linkage studies lend themselves to multicentre collaboration (as they are robust to allelic heterogeneity), much larger linkage studies of complex disease could be used in the future to identify genetic variants with moderate genetic effects. Linkage signals from such studies will provide a sure platform for subsequent association studies in the same study populations.

If we assume that variants with larger effects contribute more to the extremes of the disease or trait distributions, then study designs which seek to study extreme phenotypes may be a promising approach. Concordant (or discordant) sib pairs showing trait values (or disease age-of-onset) close to the extremes of the distribution are studied in this approach (Gu and Rao, 2001). These strategies typically require a very large sampling frame from which to select extremes. For example, Fullerton *et al.* (2003) selected the most extreme discordant and concordant sibling pairs from 34 580 sibling pairs rated for the personality trait of neuroticism, which identified a number of promising linkage peaks. Existing large-scale observational or intervention studies with banked DNA and suitable consent or large scale national biobank studies may provide alternative frameworks for such an approach.

Allele frequency of disease susceptibility variants: implications for study strategy

Although most variant sites in the population are rare, about 90% of heterozygous sites within a single individual are due to common variants (frequency >0.01) (Altshuler *et al.*, 2005). These variants may contribute significantly to genetic risk for common diseases (the common disease common variant (CDCV) hypothesis) (Reich and Lander, 2001). To the extent that this is true, a catalogue of common variants could be built to test for association with complex disease. These associations should be potentially detectable in large case-control genetic association studies.

However, the majority of variant sites in the genome are at lower frequencies, in part because deleterious variants may be subject to selection. If the genetic component of late-onset diseases is principally due to very large numbers of rare variants in numerous genes (the common disease/rare variant (CDRV) hypothesis) then the above strategy could have limited success due to inadequate power. There are good reasons to expect that a rare variant–common disease model may be more typical of complex disease than the CDCV model. The possibility that a significant fraction of the genetic variation in complex traits is due to rare alleles maintained by mutation–selection balance is supported by 25 years of research in animal genetics and by review of existing data and models (Pritchard, 2001; Wright *et al.*, 2003). Furthermore, the subset of non-synonymous coding SNPs (which are predicted on structural grounds to be deleterious) are found at significantly lower frequencies than non-synonymous and synonymous variants that are not predicted to affect

function (Fay *et al.*, 2001; Sunyaev *et al.*, 2001) suggesting that they are subject to selection.

The identification of rare variants contributing to common diseases poses a major challenge using currently available technologies. Firstly, both association and linkage approaches have low power to detect these variants, hence the need for (very) large sample sizes. This is further compounded in association studies by poor frequency matching of SNPs to causal rare variants, which again greatly reduces the power of association approaches based on SNPs (Zondervan and Cardon, 2004). Furthermore, a recent mutational origin for rare variants means that many will be superimposed on ancient core haplotypes. Multiple variants will tend to arise on the commonest haplotypes so linkage disequilibrium mapping will be less efficient.

What approaches are possible to detect rare variants?

Generally, large sample sizes and strategies favouring ascertainment of variants of larger effect will be helpful (see above). Rare variants with modest/large effect sizes should be detected with sufficiently large linkage studies, probably involving international collaborations. Genetic association studies utilising resequencing in candidate genes to identify rare variants within the study population may be possible on a large scale in the future since resequencing is technically easier than de novo sequencing (because a target sequence for comparison always exists) (Gibbs, 2005). The first step would be to catalog functional genetic variation in a large and ideally representative sample from the general population in order to permit identification of (very) rare variants. Very large sample sizes of cases will be required necessitating nationally or internationally coordinated recruitment efforts.

Construction of extended haplotypes with frequencies matched to those of the rare variants may be more feasible with family-based studies, which allow more accurate haplotype construction. Rare genetic variants may reach higher frequencies due to genetic drift in founder or genetic isolate populations and thus may be detectable only in these settings. Studies are now being established in many of these populations worldwide and this may lead to the identification of trait susceptibility variants which are rare in outbred populations. In addition, genes influencing rare recessive phenotypes should be more readily identified in inbred populations by homozygosity mapping, in which shared genomic regions that are identical-by-descent are identified by mapping (Lander and Botstein, 1987; Chapter 4, this volume), which may be the method of choice for detecting rare recessives. Finally, admixture mapping may be effective in detecting rare variants in populations in which this approach can be adopted (McKeigue, 2005).

Number of genetic variants contributing to genetic etiology: implications for study strategy

Many late-onset diseases result from the cumulative breakdown of numerous quantitatively varying physiological systems. For example, high blood pressure can follow from abnormalities in cardiac output, blood vessel architecture, renal function, fat distribution, endothelial function and central nervous system integration, or failure of diverse homeostatic mechanisms, including baroreceptors, natriuretic peptides, renin-angiotensin-aldosterone, kinin-kallikrein, adrenergic receptors and local vasodilator mechanisms – each under varying degrees of independent genetic control.

If common diseases are truly polygenic and determined by variation at very many loci, most individual effects will be too small and peripheral to the disease mechanism to be useful. The genetic (locus) complexity depends firstly on the physiological complexity of the trait. Hirschsprung disease, for example, is a rare and highly specific lack of enteric ganglia, in which eight causal loci have been identified but much of the disease variance is due to three interacting loci. In contrast,

coronary artery disease is a highly complex multidimensional trait with an exponentially increasing lifetime risk which is influenced by more than 280 risk factors (Wright *et al.*, 2002). Secondly, the distribution of gene effects is clearly important, since if most of the variance is determined by a handful of potentially detectable genes, the total number is irrelevant.

Animal models and studies of the effect of inbreeding in humans (Rudan *et al.*, 2003) suggest that the genetic basis of many complex diseases and traits is polygenic (rather than oligogenic). An estimate of the number of genetic factors underlying complex disease can be obtained by considering the fall in disease risk with decreasing degree of relationship to index cases (Risch, 1990; Schliekelman and Slatkin, 2002). A more rapid fall in risk indicates the involvement of a larger number of genetic variants. A truly polygenic model with large numbers of genetic variants each responsible for a small proportion of the phenotypic variance of a trait would require a greatly increased sample size. The implications for study design are similar to those discussed for variants which are rare or of small effect size above.

Genetic heterogeneity: implications for study strategy

Experience from published studies of Mendelian disorders and monogenic forms of complex disease suggest that extreme locus and allelic heterogeneity is the rule (Wright *et al.*, 2003). However, it is possible that experience from these conditions may not give a true picture of genetic complexity in most cases of complex disorders. Allelic diversity is greatest in high-penetrance severe dominant conditions since mutations are not expected to persist long in the population. Low-penetrance dominant or recessive mutations experience less selective pressure, tend to persist longer and may show more (due to mutation) or less (due to drift) disease allele heterogeneity. Weiss and Terwilliger (2000) suggested that late-onset disorders may show increased diversity since selection pressure is weak. However, there are few data on locus and allelic complexity in common disorders, although most published examples show substantial heterogeneity. Other potential levels of complexity include heterogeneity in expression patterns of different alleles in the same gene, epistatic and gene–environment interaction effects and protein network interactions within homeostatic and pathophysiological pathways.

Substantial allelic heterogeneity considerably reduces the power of association methods, and may make ''haplotype tagging'' (Altshuler *et al.*, 2005) less effective than linkage, admixture mapping or family-based association methods, although here again the power may also be reduced (Slager *et al.*, 2000).

Characteristics of populations in determining suitability for genetic studies

As gene discovery efforts focus on whole populations and with the increasing emphasis on association (linkage disequilibrium (LD) mapping) and other approaches such as admixture mapping, it is very important to consider the characteristics of the populations under study (Wright *et al.*, 1999). Characteristics which may be important in the design and interpretation of studies include:

- the pattern and extent of background LD that occurs in the population (large segments of LD may favour whole genome scanning but limit the ability to carry out fine mapping)
- the frequencies of alleles influencing the trait
- measures of genetic diversity (small isolated populations may have reduced genetic diversity due to few founders or drift)
- the prevalence of disease under investigation
- measures of genetic admixture
- measures of population stratification, since even low-level differentiation (e.g. between most European populations) can lead to false positive results and failure to detect genuine associations in large-scale studies (Marchini *et al.*, 2004).

- measures of variance in exposure to environmental factors (populations with limited environmental variance are favourable for studying genetic variants)
- levels of inbreeding (inbred populations are favourable for studying recessive traits)
- data on the time and circumstances of population founding events
- data on historical population sizes and bottleneck events

Other factors which may prove favourable for the support of genetic epidemiological studies are:

- the presence of good genealogical records
- the stability of the study population. Low levels of immigration and emigration result in more intact families for study.
- access to good quality medical records

Special populations

Inbred founder populations

Most complex human disease alleles identified to date are partially recessive as predicted from results in other organisms (Wright *et al.*, 2003). Large inbred pedigrees and isolated endogamous populations may be especially advantageous, because of their environmental uniformity and increased power to detect identical-by-descent (IBD) segments containing susceptibility variants (Peltonen *et al.*, 2000). These factors facilitate the ascertainment of familial disease resulting from homozygous variants of moderate to large effect and the strategy is robust in the face of allelic heterogeneity. Detection of shared homozygous identical-by-descent segments (homozygosity mapping) may also assist the detection of such recessive alleles even in common disorders. In addition, variants which are rare in outbred populations may drift to higher frequencies and be detectable in these populations (as discussed above). Many disease loci identified in this way may be ''local'' to specific populations and families.

The disadvantages of these populations include typically small population sizes, yielding fewer patients with disease and more stochastic variation. The potential difficulty in replicating positive association or linkage findings due to rare variants which may be population-specific (although this should not apply to common variants). In linkage studies of quantitative traits, the relatively small sibship sizes typically found in the most recent generations may reduce the power to identify genetic variants of small or moderate effect. Populations where polygamy is practised or high fertility is present may be particularly advantageous for such studies.

Admixed populations

Admixture mapping takes advantage of LD arising from admixture between founding populations in whom genetic variants underlying disease or trait susceptibility differ substantially in frequency (McKeigue, 2005; Chapters 4 and 5, this volume). A step beyond this is when risk alleles have become differentially fixed in one population, so that variation in disease risk is determined by environmental exposure (the common disease/fixed variant hypothesis). Admixture between ethnic groups showing high and low disease prevalence, under currently similar but differing historical environmental exposures, may be the only way to identify moderate or large effect genes contributing to such differences.

Population biobanks

Large collections of DNA samples and clinical data or ''biobanks'' are being assembled in a number of countries. These are typically very large (multigenerational) cohort studies involving hundreds of thousands of participants. The main aim of these studies is usually the characterization of genetic variants (unbiased measurement of absolute, relative and attributable disease risks and assessment of gene environment interactions) rather than gene discovery. In terms of gene discovery, these studies represent inefficient designs but nevertheless could form the platform for the

establishment of an intensively phenotyped cohort study (see above) or the recruitment of very large series of affected sib pairs, trios or quartets. They also provide a very large sampling frame from which to select and study extreme discordant or concordant sib pairs.

Family-based and population-based strategies

Family-based approaches

Linkage analysis has proved to be very effective for identifying genes concerned with rare Mendelian disorders and Mendelian forms of genetically complex ones. Thus, it is the major strategy used for identification of genetic variants of large effect. However large genetic effects tend to result in selection against these alleles, resulting in rare diseases or rare forms of common diseases. Power is governed by the locus-specific sibling recurrence risk ratio for a binary trait, or the locus-specific heritability for a continuous trait. The above discussion implies that large numbers of families will have to be studied and hence international collaborations are required. The low power to detect genetic variants of small effect suggests that this approach may be much less effective for common complex diseases.

It will become increasingly important to combine pedigrees, perhaps from different countries, to acquire sufficient power for linkage-based approaches. The failure to identify high penetrance variants in important pathways suggests that effect sizes are generally modest to small. Linkage approaches have not proved successful to date in identifying such small genetic effects in complex diseases (Merikangas and Risch, 2003). However, recent pedigree-based studies carried out by deCODE Genetics in Iceland have reported promising results with denser marker scans and larger pedigrees than in past studies (Freimer and Sabatti, 2004).

This pedigree- or family-based approach has a number of other advantages:

- the ability to make use of *familial correlations* to increase power using strategies such as the EDAC (''extreme discordant and concordant'') design
- the ability to obtain *information on genetic haplotypes* (most readily determined with family data, particularly over extended genetic distances, despite recent improvements in prediction)
- the ability to carry out *transmission disequilibrium tests* (TDT) by comparing transmitted with non-transmitted parental alleles
- the ability to provide *parent-of-origin information* (e.g. the relationship between early life events and adult diseases may involve maternally or paternally imprinted loci, such as IGF2)

Population-based approaches

Association studies (LD mapping) measure differences in the frequency of genetic variants between unrelated affected individuals (cases) and controls. These studies are premised on the hypothesis that common genetic variants underlie susceptibility to common diseases. Risch and Merikangas (1996) demonstrated that linkage approaches were only likely to be feasible for genotypic relative risks greater than four. Effects of this magnitude are unlikely for most variants influencing complex diseases so that association approaches using dense marker maps (assuming high LD between marker and disease allele) should have greater power. In reality, the power of association approaches is compromised by incomplete LD between the marker (or tested candidate variant) and the disease variant and also by differences in allele frequency between the marker and disease variants (Abel and Muller-Myhsok, 1998; Zondervan and Cardon, 2004). Power can be increased by employing ''enriched'' case series, for example by selecting cases with early-onset disease or with a family history of disease, which will selectively enrich for rare variants

(Antoniou and Easton, 2003). The use of ''hypernormal'' controls, as in a comparison of hypertensives and controls with unusually low blood pressure for their age, may be more efficient than using random controls (Morton and Collins, 1998).

The role of population structure as an unmeasured confounder may have been exaggerated. In apparently homogeneous populations, hidden population stratification is unlikely to be a serious problem (Morton and Collins, 1998; Wacholder *et al.*, 2002). However in some highly stratified populations, such as African Americans where admixture varies between individuals, it is necessary to control for population structure. In the last few years, the development of statistical methods to control for population stratification within the analysis (Pritchard *et al.*, 2000) has reinstated standard case-control designs in genetic epidemiology. It has been suggested that attention should be refocused away from concerns about population stratification towards the more immediate problems of multiple testing, encouraging larger-scale studies, obtaining data on background LD levels and exploring the utility of multilocus associations (Cardon and Bell, 2001; Cardon and Palmer, 2003).

Many false positive associations have been reported due to a combination of multiple testing and publication bias (Zondervan *et al.*, 2004). Other reasons include chance, insufficient attention given to standard procedures for identifying false positives, and failure to adopt traditional epidemiological methods to limit bias and confounding. A robust causal inference framework (including the requirement for independent replication of findings) is therefore necessary to interpret results (Campbell and Rudan, 2002).

Case-control designs are more efficient for a given cost, can be established more quickly and can give more attention to disease-specific exposures and outcomes. Cohort studies are valuable when a valid measure of exposure requires its measurement before disease onset or at multiple time periods over several years. Cohorts also provide for identical exposures in cases and controls and are less prone to selection bias. For example, they do not exclude cases who die soon after diagnosis, and they can document data on multiple outcomes (Colhoun *et al.*, 2003).

These approaches can be extended by recruiting offspring to facilitate assessment of haplotypes or by recruitment of parents, particularly in studies of childhood disorders. This allows the use of family-based association methods such as FBAT (Lange *et al.*, 2002) or examination of parent-of-origin effects. Consideration needs to be given to whether the additional recruitment effort and genotyping expense is justified in each particular study setting.

Direct association

In direct association studies, genetic variants are studied in selected genes which are believed to cause disease. This approach could proceed by cataloging and making publically available all common variants in coding and regulatory regions of genes (Carlson *et al.*, 2004; Altshuler *et al.*, 2005) and then comparing frequencies of these variants in cases and controls. However, only a small proportion of human genes have been resequenced in sufficient numbers of individuals so that in-depth characterisation of genetic variation is available. In the near term, this strategy can exploit extensive information from the HapMap project on common variants within candidate genes and surrounding regions that are evolutionarily conserved (Altshuler *et al.*, 2005). These studies are most meaningful when applied to functionally significant variants with a clear biological relation to disease mechanisms or the trait under study. There are no widely accepted criteria for the selection of candidate variants. Variants that represent credible candidates include: ones associated with familial forms of disease; in genes encoding proteins involved in disease pathogenesis; and in genes encoding xenobiotic enzymes thought to interact with environmental exposures.

In practice, it is necessary to decide on a set of candidate genes likely to be associated with a condition, to identify the SNPs likely to be most

informative in the study population and to use them for association studies in large well-characterized groups of patients and controls. Of key importance is that such studies have adequate power to detect or exclude genotype risk ratios of a moderate size. In future, a substantial fall in the cost of resequencing may make this affordable so that rare variants can be evaluated in a genetic association framework.

Indirect association

Indirect association studies investigate the association resulting from LD between the genetic marker and the true disease susceptibility variant(s). This identifies segments of the genome that are inherited from a common ancestor. The ongoing HapMap project is designed to facilitate such studies and to identify minimum sets of ''tagging SNPs'' that show strong associations with known common variants (Altshuler *et al.*, 2005). Uncertainties arise because shared segment sizes are extremely stochastic resulting in variable extents of LD even between very closely associated markers. LD is affected by many processes, including mutation, gene conversion, recombination, selection and population history. This results in the non-linear decay of LD with time, depending on the evolutionary history of each variant, so that stronger disease associations may be present with variants at a greater distance from the disease variant.

LD mapping is powerful when cases share the same allele IBD from a common ancestor, as in rare Mendelian disorders. However, the power of this approach is dependent on the number of founder chromosomes in the study population and thus power falls when there is locus and allelic heterogeneity. For complex diseases, the number of founder alleles may only be reduced in isolate populations with very few founders or with recent population bottleneck events.

There is currently limited experience with high density SNP genotyping (''whole genome association'') studies although these are now affordable and many such studies will take place in the coming years. Current problems include the need to define the best SNPs to use for either scoring functional variants or tagging haplotypes in each gene (Altshuler *et al.*, 2005); the need to gain a better understanding of patterns of LD and the extent to which they are locus-specific and population-specific. A sequence-based approach for selecting SNPs based on likely functional effects may provide power advantages by reducing the number of SNPs and significance level required (Botstein and Risch, 2003) and may allow detection of associations with alleles at lower frequency. In the short term, whole genome association mapping may be most successful in isolate populations with a small effective number of unrelated founders and with demonstrably large areas of LD throughout the genome (Service *et al.*, 2006). As the density of markers typed rises with falling costs to >1 million markers, then this approach will merge with the direct association approach involving a systematic evaluation of candidate gene variants. Populations, such as those from Africa, which exhibit low LD may be favored in the future as genotyping costs continue to fall because they will permit greater precision in identifying causal variants and will be less subject to confounding due to LD.

An important problem with the association mapping approach is the reduced power due to incomplete LD between marker and disease susceptibility alleles which has resulted in greatly underestimating study sample sizes. It is more efficient to increase sample size than to increase the density of markers (Long and Langley, 1999) since the relative sample size increases only linearly with the logarithm of the number of comparisons made (Witte *et al.*, 2000). Current power calculations for high density SNP genotyping studies are premised on a multistage approach (van den Oord and Sullivan, 2003; Satagopan *et al.*, 2004). The aim is to increase genotyping efficiency by screening all markers on a subset of participants in stage 1 and by only using the most significant

markers in a larger group of participants in subsequent stages.

Approaches to the analysis of these studies are evolving. Two stage designs have been proposed in which a proportion (typically ≥0.3) of the samples are genotyped for the full genetic marker set in a first stage, and a proportion of the markers are genotyped in the remaining samples in a second stage. Data from both stages are then analyzed jointly to maximize study power (Skol *et al.*, 2006). Selection of markers to test in Stage 2 requires a ranking procedure which will have a high probability of including variants which are true positives (Zaykin and Zhivotovsky, 2005). Control of false positives due to multiple testing requires permutation testing, use of Bayesian methods or some other equivalent technique.

Journal articles looking at observed versus expected distributions of test statistics in these studies have shown inflation of test statistics and have pointed to the importance of differential bias in genotype scoring, between cases and controls, and population stratification. This underscores the importance of study design in the recruitment and testing of cases and controls and the need to estimate and correct for confounding due to population stratification (Clayton *et al.*, 2005; Hirschorn and Daly, 2005; Wang *et al.*, 2005). Interpretation of positive findings after controlling for confounding factors and multiple test correction will require replication in independent datasets and attention to established causal criteria (see below).

Integrated study designs, with simultaneous collection of family-based and population-based controls have been proposed (e.g. two-stage case-family, case-family-control and case-control family designs) (Zhao *et al.*, 1999). These can be analyzed in stages, or combined in a single analysis that estimates within-family and between-family associations simultaneously. Alternatively, collecting both population and family controls for the same cases means that positive findings in the case-control design can be independently verified in the family design. Analytic methods that combine information from both linkage and association (LD) approaches, such as those measuring both IBD and allele-specific IBD sharing and those detecting associations in large pedigrees, may have advantages (Whittemore and Nelson, 1999).

Expression quantitative trait loci (QTL)

Human QTL studies aim to understand the molecular mechanisms underlying individual variation in biological quantitative traits, many of which are established risk factors for complex common diseases (de Koning and Haley, 2005). Traditional approaches to QTL mapping in humans have largely focused on the identification of loci affecting just one or a few biological quantitative traits (Kendziorski *et al.*, 2006). However, current microarray technologies allow measurement of many thousand gene expression levels each of which are also quantitative traits that have been shown to be highly variable and heritable (Cheung *et al.*, 2005). A focus of interest in QTL mapping is now shifting to a more systematic and comprehensive characterization of quantitative biological variation from the perspective of the genome itself. The exponential model of QTL effect sizes (based on Fisher's infinitesimal model of adaptive evolution) predicts that it should be possible to explain a substantial amount of quantitative genetic variation by a limited number of genes with large effect sizes. The variants that affect gene expression (expression QTL, eQTL) are thereby expected to have an important role in generating quantitative biological variation (Farrall, 2004).

The design of studies to identify these eQTL is similar to traditional QTL studies in that they aim to identify the locations in the genome to which the expression traits are linked (Kendziorski *et al.*, 2006). Early experiments that combined expression profiling with classical genetic mapping approaches, such as the studies of Cheung *et al.* (2005), have revealed a wealth of ''expression phenotypes'' in the human genome. These eQTL studies which investigate thousands of expression

transcripts face significant challenges in statistical inference (Kendziorski *et al.*, 2006) and the need for new computational resources to visualize and explore data from combined genome-wide expression and linkage studies (Carlborg *et al.*, 2005; Mueller *et al.*, 2006).

Interaction effects

Gene–gene interactions (epistasis)

Molecular interactions are an essential feature of most homeostatic developmental and pathophysiological pathways at the cellular level and epistasis has been shown to be important for determining phenotypes for a number of rare Mendelian diseases.

It has been suggested that efficient homeostasis may mean that extreme phenotypes due to single mutations are rare and that genetic mutations affecting multiple pathways are required before extreme values of biological traits are found. It has thus been argued that epistatic interactions are an important component of the architecture of most common traits (Templeton *et al.*, 2000). There is considerable biological and empirical evidence from human studies for epistasis in complex traits (Marchini *et al.*, 2005). Variants with large marginal effects relative to their interaction effects will be detected by the current one-locus association approach. However, variants with significant interaction effects but less marked (marginal) single locus effects may be better detected by two or three locus approaches (Marchini *et al.*, 2005).

Epistasis may represent another dimension of complexity that is to date unaccounted for in the large majority of studies, and thus could be one of the reasons for the failure to replicate so many reported disease associations. If interactive effects are equal to or larger than marginal effects of genetic variants then failure to evaluate 2- or 3-locus models may result in important genetic variants remaining undiscovered. However, these will be challenging to study due to the need for very large sample sizes and computing power. In addition, apparent statistical epistatic effects can be an artefact due to the scale used to model interactions (Carlborg and Haley, 2004) and are not always biologically relevant (Clayton and McKeigue, 2001).

Gene–environment interactions

It has been argued that the ability to model the joint effects of genotype and environmental exposure does not in general make it easier to detect genes or environmental factors that influence disease risk (Clayton and McKeigue, 2001).

Most environmental variables relevant to common diseases are either not measurable or not reliably measurable, even with the best study designs. The focus of gene–environment interaction studies will therefore be on a restricted subset of exposures that are both reliably measured and potentially modifiable and hence of potential public health importance. The recent changes in prevalence of many common diseases suggests that such modifiable influences (e.g. diet, lifestyle) may be major contributors.

Mendelian randomization

Mendelian randomization (MR) refers to the random assortment of genes from parents to offspring that occurs during gamete formation and conception. It has been suggested that this natural randomization process, which occurs at conception to determine a person's genotype, could be used to assess the causal nature of some established associations between environmental exposures and disease occurrence (Davey Smith and Ebrahim, 2003). The key assumption is that, combining the information derived from genotype–disease and genotype–phenotype studies, taking into account MR should make it possible to estimate the association between a modifiable phenotype (e.g. an established environmental risk factor) and disease that is free of the confounding

and reverse causation typical of classical epidemiological studies (Minelli *et al.*, 2004). Thus, this approach could potentially contribute to causal inference in established associations between modifiable environmental risk factors and common complex diseases (Smith and Ebrahim, 2004; Davey Smith *et al.*, 2005). The genetically determined (lifelong) ''exposure'' may however not be biologically equivalent to the measured environmental exposure. Plotting the log odds ratio of genotype–disease association against the mean genotype–phenotype difference may identify significant departures from the assumptions underlying MR. Testing for differences between studies that only report genotype–phenotype or genotype–disease association and those reporting on both, may help to detect bias in the results of the former (Minelli *et al.*, 2004).

Although attractive in theory, the utility of this approach in practice remains to be demonstrated and may be restricted by a number of important limitations including confounding by polymorphisms that are in linkage disequilibrium with the polymorphism under study; the large number of phenotypes associated with disease; the lack of suitable polymorphisms for studying phenotypes of interest; and the buffering of the effects of genetic variation during development (canalization).

Interpretation of results: causal inference

Problems with the correct interpretation of genetic association results has drawn attention to the importance of having an agreed framework for causal inference (Colhoun *et al.*, 2003; Campbell and Rudan, 2002). Several criteria may be helpful in interpreting the results of genetic association studies.

Experimental evidence

Early genetic linkage studies required corroboration by cytogenetic or somatic cell hybrid studies before being accepted. There is a need to combine genetic/biological and epidemiological/statistical data with functional studies, such as those using knock-out animals, cell lines, gene expression or enzyme studies, in order to complement observational data from genetic epidemiology studies. Traditional epidemiology has been limited in the extent to which it can call on experimental data to interpret observational data since randomized controlled trials are only ethical and feasible to implement for a minority of research questions. The above genetic approaches coupled to MR study designs (Davey Smith and Ebrahim, 2003) greatly extend the opportunities to check genetic associations by use of experimental data and to reach more robust conclusions. However, this will require close multidisciplinary collaboration between epidemiology and biology research groups. The discovery and confirmation of the role of the human mismatch repair genes (*hMLH1* and *hMSH2*) in colorectal cancer provides a good example of how observational epidemiological data and experimental biological data can combine to provide robust proof (Campbell and Rudan, 2002; Chapter 17, this volume).

Consistency of association

Attempting to repeat the finding of a positive association at least once in an independent sample, ideally in the same population, using either a split sample or repeated sample approach is essential. Confirmation using a family-based TDT approach, when the initial report was from a case-control comparison, would demonstrate both linkage and association and strengthen the evidence for a causal association. Repeatability (the probability that a second association study is also positive in the same population) varies with sample size and may be low in sample sizes of the order of 100–500 cases (Long and Langley, 1999). The sample size of the replication study should be at least as large as the original study and sample size calculations should be based on

a more conservative effect size than that reported in the original study (Colhoun *et al.*, 2003). Experience suggests that a finding of $p < 0.01$ in two follow-up studies or $p < 0.001$ in a single follow-up study are each strongly predictive of future replication (Lohmueller *et al.*, 2003). If replication is considered essential before a positive result is accepted as true, then a moderate false positive rate is acceptable (Thomas and Clayton, 2004) and preferable to a very conservative strategy, which will tend to result in false negatives. This has been formalized in the multistage strategy currently advocated for whole genome association studies (Hirschhorn and Daly, 2005). A first threshold is set to maintain power to detect variants that account for only a small percentage of the phenotypic variance and so that negatives can safely be discarded. Following this, a second independent sample is analysed with permutation testing to define statistical significance thresholds.

Where LD provides the basis for the observed association, it is not likely to be consistent across populations, since LD depends on population history. Association may be present with one allele in one population, another allele in a second population and no association found with either of these alleles in a third population. Even when the causal variant is under investigation, a genetic variant may be more or less important in different populations depending, for example, on its allele frequencies in those populations.

Biological plausibility

The nature of the genetic variant which has been shown to be associated with a disease may provide clues to interpretation by classifying the variant according to biological effect:

Possible interpretation of different types of variant

- *pathogenic*: frameshift; nonsense; splice variant
- *possibly pathogenic*: non-conservative amino acid substitution
- *probable polymorphism*: conservative amino acid substitution; present in controls
- *definite polymorphism*: synonymous variant present in controls with no predicted effect on RNA (ribonucleic acid) splicing.

Variants can be prioritized by a variety of bioinformatic techniques, for example according to evolutionary conservation, effect on protein structure and potential effects on RNA splicing (Ng and Henikoff, 2002; Wang and Moult, 2001; Ramensky *et al.*, 2002). As the typed variant may just be in LD with the variant that directly contributes to the disease, it is also necessary to prioritize the immediate genomic region by defining the relevant haplotypes and tagging SNPs using HapMap data (Altshuler *et al.*, 2005). Analysis of other properties of the associated region, such as evolutionary conservation, the distribution of neighboring genes and predicted micro RNAs, may be important. The genes can then be prioritized using a wide variety of criteria, such as expression changes in disease, co-expression/homology/functional similarity with known disease-associated genes and position with respect to known susceptibility/associated regions (Botstein and Risch, 2003).

Strength of association

In traditional epidemiological studies, stronger associations are more likely to be causal. However, this may be problematic in studies of genetic variants. In indirect association studies, the strength of association depends on the extent of LD and the relative frequencies of marker and susceptibility alleles. Thus, strong associations may be reported for a marker in strong LD with susceptibility allele of weak effect and vice versa. Furthermore, strong association does not imply tight linkage. Strength of association is also affected by the underlying genetic model and will vary depending on whether this is dominant, recessive or multiplicative.

Interpretation of results: reporting and characterizing genetic effects

It is important that a framework for the reporting of genetic associations is adopted internationally so that communication of results and meta-analysis of studies is facilitated (Little *et al.*, 2002). Key questions to be addressed after the identification of a genetic variant's role in a disease process or a disease-associated trait are:

- what is the prevalence of the gene variant(s) in different populations?
- what is the risk (relative and attributable) of disease associated with these variants?
- what is the contribution of particular genetic variants to the occurrence of disease in different populations? (population attributable risk)
- what is the risk of disease associated with gene–gene and gene–environment interactions?

The frequency of genetic variants can be estimated from large unselected cohorts or large randomly sampled control series in different populations. Thus, although selected groups may be utilized to increase study power in gene discovery studies (Little *et al.*, 2002), gene characterization studies require representative population samples in order to make unbiased estimates of risk. Absolute risks can be measured in population-based cohort studies and relative risks in cohort or genetic association studies. Attributable risk measures the absolute size of the excess risk associated with the variant. It thus expresses the size of excess risk in individuals with particular genotypes. The large national biobank studies noted above will have sufficient power to yield very precise estimates of these parameters.

International collaboration

The translation of the knowledge gained through genetic and genomic research into clinical medicine and public health benefits will first require studies of much larger scale than most of those currently being undertaken. In order to make full use of the wealth of information that is being gathered across the world, a massive and well-coordinated initiative will be needed to bring together genetic epidemiological research on a global scale (Khoury, 2004). Building on disease-specific collaborations which have assembled data to perform meta-analyses and encouraging collaboration between large existing population-based collections would be an important first step. Developing approaches to pool the information across the many data collections and to integrate the data collected under different study designs will become necessary, and early efforts are already underway e.g. the P3G (Public Population Project in Genomics) project. To collate and integrate data on human genes and the impact of human genome variation on population health, the Human Genome Epidemiology Network was launched (HuGENet) (Khoury, 2004).

Adoption of uniform protocols and procedures internationally would facilitate collaboration and integration of data across national boundaries. This would include documentation and clarification of the variation which exists in national laws and regulations with respect to obtaining, storing and using the biological data and information, which could serve as an obstacle to any sharing of data across national borders or render it unlawful (Kaye, 2006). Progress will require wide cooperation across many disciplines and boundaries and the active involvement of many segments of society for the benefit of its citizens.

Conclusions

The primary goals in mapping complex disease traits are not to understand disease in the majority of cases, which the biological complexity suggests may be unrealistic in the short- to medium-term, but to identify biological mechanisms or pathways which may lead to novel preventive measures or drug targets. The priority is therefore to identify large-to-moderate rather than average genetic effects using the most powerful strategies available. A sober consideration of

what is known about the genetic architecture of complex disease is likely to lead to improved study design and ultimately more successful investment in this research field. Genetic epidemiological studies will increasingly be characterized by international collaboration, large sample sizes and more determined attempts to utilize the uniqueness of population genetic characteristics found in diverse world populations. In addition, multidisciplinary teams employing joint epidemiological and statistics as well as biological and informatics expertise to understand increasingly complex patterns of genetic, biological and medical data.

REFERENCES

Abel, L. and Muller-Myhsok, B. (1998). Maximum-likelihood expression of the transmission/disequilibrium test and power considerations. *Am J Hum Genet*, **63**, 664–7.

Altshuler, D., Brooks, L.D., Chakravarti, A. *et al.* International HapMap Consortium. (2005). A haplotype map of the human genome. *Nature*, **437**, 1299–320.

Antoniou, A.C. and Easton, D.F. (2003). Polygenic inheritance of breast cancer: Implications for design of association studies. *Genet Epidemiol*, **25**, 190–202.

Barton, N.H. and Keightley, P.D. (2002). Understanding quantitative genetic variation. *Nature Rev Genet*, **3**, 11–21.

Botstein, D. and Risch, N. (2003). Discovering genotypes underlying human phenotypes: past successes for Mendelian disease, future approaches for complex disease. *Nat Genet*, **33** Suppl, 228–37.

Campbell, H. and Rudan, I. (2002). Interpretation of genetic association studies in complex disease. *Pharmacogenomics J*, **2**, 349–60.

Carlborg, O. and Haley, C.S. (2004). Epistasis: too often neglected in complex trait studies? *Nat Rev Genet*, **5**, 618–25.

Cardon, L.R. and Bell, J.I. (2001). Association study designs for complex diseases. *Nat Rev Genet*, **2**, 91–9.

Cardon, L.R. and Palmer, L.J. (2003). Population stratification and spurious allelic association. *Lancet*, **361**, 598–604.

Carlborg, O., De Koning, D.J., Manly, K.F. *et al.* (2005). Methodological aspects of the genetic dissection of gene expression. *Bioinformatics*, **21**, 2383–93.

Carlson, C.S., Eberle, M.A., Kruglyak, L. and Nickerson, D.A. (2004). Mapping complex disease loci in whole-genome association studies. *Nature*, **429**, 446–52.

Cheung, V.G., Spielman, R.S., Ewens, K.G. *et al.* (2005). Mapping determinants of human gene expression by regional and genome-wide association. *Nat Genet*, **437**, 1365–9.

Claus, E.B., Risch, N.J. and Thompson, W.D. (1990). Using age of onset to distinguish between subforms of breast cancer. *Ann Hum Genet*, **54**, 169–77.

Clayton, D. and McKeigue, P.M. (2001). Epidemiological methods for studying genes and environmental factors in complex diseases. *Lancet*, **358**, 1356–60.

Clayton, D.G., Walker, N.M., Smyth, D.J. *et al.* (2005). Population structure, differential bias and genomic control in a large-scale, case-control association study. *Nat Genet*, **37**, 1243–6.

Cohen, J.C., Kiss, R.S., Pertsemlidis, A. *et al.* (2004). Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science*, **305**, 869–72.

Colhoun, H.M., McKeigue, P.M. and Davey Smith, G. (2003). Problems of reporting genetic associations with complex outcomes. *Lancet*, **361**, 865–72.

Davey Smith, G. and Ebrahim, S. (2003). 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*, **32**, 1–22.

Davey Smith, G., Ebrahim, S., Lewis, S. *et al.* (2005). Genetic epidemiology and public health: hope, hype, and future prospects. *Lancet*, **366**, 1484–98.

de Koning, D.J. and Haley, C.S. (2005). Genetical genomics in humans and model organisms. *Trends Genet*, **21**, 377–81.

Devlin, B., Bacanu, S.A. and Roeder, K. (2004). Genomic control to the extreme. *Nat Genet*, **36**, 1129–30.

Farrall, M. (2004). Quantitative genetic variation: a post-modern view. *Hum Mol Genet*, **13** Spec No 1, R1–7.

Fay, J.C., Wyckoffand, G.J. and Wu, C.I. (2001). Positive and negative selection on the human genome. *Genetics*, **158**, 1227–34.

Freimer, N. and Sabatti, C. (2004). The use of pedigree, sib-pair and association studies of common diseases for genetic mapping and epidemiology. *Nat Genet*, **36**, 1045–51.

Fullerton, J., Cubin, M., Tiwari, H. *et al.* (2003). Linkage analysis of extremely discordant and concordant sibling pairs identifies quantitative-trait loci that influence variation in the human personality trait neuroticism. *Am J Hum Genet*, **72**, 879–90.

Gibbs, R. (2005). Deeper into the genome. *Nature*, **437**, 1233–4.

Gu, C. and Rao, D.C. (2001). Optimum study designs. *Adv Genet*, **42**, 439–57.

Hirschhorn, J.N. and Daly, M.J. (2005). Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet*, **6**, 95–108.

Kaye, J. (2006). Do we need a uniform regulatory system for biobanks across Europe? *Eur J Hum Genet*, **14**, 245–8.

Keavney, B. (2002). Genetic epidemiological studies of coronary heart disease. *Int J Epidemiol*, **31**, 730–6.

Kendziorski, C.M., Chen, M., Yuan, M., Lan, H. and Attie, A.D. (2006). Statistical methods for expression quantitative trait loci (eQTL) mapping. *Biometrics*, **62**, 19–27.

Khoury, M.J. (2004). The case for a global human genome epidemiology initiative. *Nat Genet*, **36**, 1027–8.

Lander, E.S. and Botstein, D. (1987). Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. *Science*, **236**, 1567–70.

Lange, C., DeMeo, D.L. and Laird, N.M. (2002). Power and design considerations for a general class of family-based association tests: quantitative traits. *Am J Hum Genet*, **71**, 1330–41.

Little, J., Bradley, L., Bray, M.S. *et al.* (2002). Reporting, appraising, and integrating data on genotype prevalence and gene–disease associations. *Am J Epidemiol*, **156**, 300–10.

Lohmueller, K.E., Pearce, C.L., Pike, M., Lander, E.S. and Hirschhorn, J.N. (2003). Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet*, **33**, 177–82.

Long, A.D. and Langley, C.H. (1999). The power of association studies to detect the contribution of candidate genetic loci to variation in complex traits. *Genome Res*, **9**, 720–31.

Lyons, M.J., Eisen, S.A., Goldberg, J. *et al.* (1998). A registry-based twin study of depression in men. *Arch Gen Psychiat*, **55**, 468–72.

Macgregor, S., Knott, S.A., White, I. and Visscher, P.M. (2003). Longitudinal variance-components analysis of the Framingham Heart Study data. *BMC Genet*, **4** Suppl 1, S22.

Marenberg, M.E., Risch, N., Berkman, L.F., Floderus, B. and de Faire, U. (1994). Genetic susceptibility to death from coronary heart disease in a study of twins. *N Engl J Med*, **330**, 1041–6.

Marchini, J., Cardon, L.R., Phillips, M.S. and Donnelly, P. (2004). The effects of human population structure on large genetic association studies. *Nat Genet*, **36**, 512–17.

Marchini, J., Donnelly, P. and Cardon, L.R. (2005). Genome-wide strategies for detecting multiple loci that influence complex diseases. *Nat Genet*, **37**, 413–17.

Merikangas, K.R. and Risch, N. (2003). Genomic priorities and public health. *Science*, **302**, 599–601.

McKeigue, P.M. (2005). Prospects for admixture mapping of complex traits. *Am J Hum Genet*, **76**, 1–7.

Millar, J.K., Pickard, B.S., Mackie, S. *et al.* (2005). DISC1 and PDE4B are interacting genetic factors in schizophrenia that regulate cAMP signaling. *Science*, **310**, 1187–91.

Minelli, C., Thompson, J.R., Tobin, M.D. and Abrams, K.R. (2004). An integrated approach to the meta-analysis of genetic association studies using Mendelian randomization. *Am J Epidemiol*, **160**, 445–52.

Morton, N.E. and Collins, A. (1998). Tests and estimates of allelic association in complex inheritance. *Proc Natl Acad Sci USA*, **95**, 11389–93.

Mueller, M., Goel, A., Thimma, M. *et al.* (2006). eQTL Explorer: integrated mining of combined genetic linkage and expression experiments. *Bioinformatics*, **22**, 509–11.

Ng, P.C. and Henikoff, S. (2002). Accounting for human polymorphisms predicted to affect protein function. *Genome Res*, **12**, 436–46.

Peltonen, L., Palotie, A. and Lange, K. (2000). Use of population isolates for mapping complex traits. *Nat Rev Genet*, **1**, 182–90.

Pharoah, P.D., Antoniou, A., Bobrow, M. *et al.* (2002). Polygenic susceptibility to breast cancer and implications for prevention. *Nat Genet*, **31**, 33–6.

Pritchard, J.K., Stephens, M., Rosenberg, N.A. and Donnelly, P. (2000). Association mapping in structured populations. *Am J Hum Genet*, **67**, 170–81.

Pritchard, J.K. (2001). Are rare variants responsible for susceptibility to complex diseases? *Am J Hum Genet*, **69**, 124–37.

Ramensky, V., Bork, P. and Sunyaev, S. (2002). Human non-synonymous SNPs: server and survey. *Nucl Acids Res*, **30**, 3894–900.

Reich, D.E. and Lander, E.S. (2001). On the allelic spectrum of human disease. *Trends Genet*, **17**, 502–10.

Risch, N. (1990). Linkage strategies for genetically complex traits. I. Multilocus models. *Am J Hum Genet*, **46**, 222–8.

Risch, N. and Merikangas, K. (1996). The future of genetic studies of complex human diseases. *Science*, **273**, 1516–17.

Rudan, I., Campbell, H., Carothers, A. *et al.* (2003). Inbreeding and the genetic complexity of human essential hypertension. *Genetics*, **163**, 1011–21.

Satagopan, J. M., Venkatraman, E. S. and Begg, C. B. (2004). Two-stage designs for gene-disease association studies with sample size constraints. *Biometrics*, **60**, 589–97.

Schliekelman, P. and Slatkin, M. (2002). Multiplex relative risk and estimation of the number of loci underlying an inherited disease. *Am J Hum Genet*, **71**, 1369–85.

Service, S., DeYoung, J., Karayiorgou, M. *et al.* (2006). Magnitude and distribution of linkage disequilibrium in population isolates and implications for genome-wide association studies. *Nat Genet*, **38**, 556–60.

Skol, A. D., Scott, L. J., Abecasis, G. R. and Boehnke, M. (2006). Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet*, **38**, 209–13.

Slager, S. L., Huang, J. and Vieland, V. J. (2000). Effect of allelic heterogeneity on the power of the transmission disequilibrium test. *Genet Epidemiol*, **18**, 143–56.

Smith, G. D. and Ebrahim, S. (2004). Mendelian randomization: prospects, potentials, and limitations. *Int J Epidemiol*, **33**, 30–42.

Sunyaev, S., Ramensky, V., Koch, I. *et al.* (2001). Prediction of deleterious human alleles. *Hum Mol Genet*, **10**, 591–7.

Templeton, A. R., Weiss, K. M., Nickerson, D. A., Boerwinkle, E. and Sing, C. F. (2000). Cladistic structure within the human lipoprotein lipase gene and its implications for phenotypic association studies. *Genetics*, **156**, 1259–75.

Thomas, D. C. and Clayton, D. G. (2004). Betting odds and genetic associations. *J Natl Cancer Inst*, **96**, 421–3.

Van den Oord, E. J. and Sullivan, P. F. (2003). False discoveries and models for gene discovery. *Trends Genet*, **19**, 537–42.

Wacholder, S., Rothman, N. and Caporaso, N. (2002). Counterpoint: bias from population stratification is not a major threat to the validity of conclusions from epidemiological studies of common polymorphisms and cancer. *Cancer Epidemiol Biomarkers Prev*, **11**, 513–20.

Wang, W. Y. S., Barrat, B. J., Clayton, D. G. and Todd, J. A. (2005). Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet*, **6**, 109–18.

Wang, Z. and Moult, J. (2001). SNPs, protein structure, and disease. *Hum Mutat*, **17**, 263–70.

Weiss, K. M. and Terwilliger, J. D. (2000). How many diseases does it take to map a gene with SNPs? *Nat Genet*, **26**, 151–7.

Whittemore, A. S. and Nelson, L. M. (1999). Study design in genetic epidemiology: theoretical and practical considerations. *J Natl Cancer Inst Monogr*, **26**, 61–9.

Witte, J. S., Elston, R. C. and Cardon, L. R. (2000). On the relative sample size required for multiple comparisons. *Stat Med*, **19**, 369–72.

Wright, A., Carothers, A. and Campbell, H. (2002). Gene environment interaction: the BioBank UK Study. *Pharmacogenomics J*, **2**, 75–82.

Wright, A., Charlesworth, B., Rudan, I., Carothers, A. and Campbell, H. (2003). A polygenic basis for late-onset disease. *Trends Genet*, **19**, 97–106.

Wright, A. F., Carothers, A. D. and Pirastu, M. (1999). Population choice in mapping genes for complex diseases. *Nat Genet*, **23**, 397–404.

Wright, A. F. and Hastie, N. D. (2001). Complex genetic diseases: controversy over the Croesus code. *Genome Biol*, **2**, 2007.1–2007.8.

Zaykin, D. V. and Zhivotovsky, L. A. (2005). Ranks of genuine associations in whole-genome scans. *Genetics*, **171**, 813–23.

Zhao, L. P., Araqaki, C., Hsu, L. *et al.* (1999). Integrated designs for gene discovery and characterization. *J Natl Cancer Inst Monogr*, **26**, 71–80.

Zondervan, K. T., Cardon, L. R. (2004). The complex interplay among factors that influence allelic association. *Nat Rev Genet*, **5**, 89–100.

7

Diseases of protein misfolding

Christopher M. Dobson

Introduction

The ability of even the most complex protein molecules to fold to their biologically functional states is perhaps the most fundamental example of biological self-assembly, one of the defining characteristics of living systems (Vendruscolo *et al.*, 2003). The process of protein folding in the cellular environment can, in principle, begin whilst the nascent chain is still attached to the ribosome, and there is clear evidence that some proteins do fold at least partially in such a co-translational manner (Hardesty and Kramer, 2001). Other proteins are known to undergo the major part of their folding only after release from the ribosome, whilst yet others fold in specific cellular compartments such as the endoplasmic reticulum (ER) or mitochondria following translocation through membranes (Hartl and Hayer-Hartl, 2002). Although the fundamental principles underlying the mechanism of folding are unlikely to differ in any significant manner from those elucidated from in vitro studies, many details of the way in which individual proteins fold will undoubtedly depend on the environment in which they are located. In particular, as incompletely folded chains inevitably expose to the outside world many regions of the polypeptide molecule that are buried in the native state, such species are prone to inappropriate interactions with other molecules within the complex and crowded cellular environment. Such interactions can result both in the disruption of normal cellular processes and in self-association or aggregation. There is evidence from in vitro studies that under some circumstances non-native interactions form transiently to bury highly aggregation prone regions such as exposed hydrophobic surfaces (Capaldi *et al.*, 2002; Hore *et al.*, 1997). To cope with this problem more generally, living systems have evolved a range of elaborate strategies to prevent interactions with other molecules prior to the completion of the folding process in vivo (Gething and Sambrook, 1992; Hartl, 2000; Hartl and Hayer-Hartl, 2002).

Of particular importance in this context is the large number of molecular chaperones that are present in all types of cells and cellular compartments, Figure 7.1. Despite their similar general role in enabling efficient folding and assembly to take place, their specific functions can differ substantially and it is evident that many types of chaperone work in tandem with each other (Hartl and Hayer-Hartl, 2002). Some molecular chaperones have been found to interact with nascent chains as they emerge from the ribosome, and bind rather non-specifically to protect aggregation-prone regions rich in hydrophobic residues. Others are involved in guiding later stages of the folding process, particularly for complex proteins including oligomeric species and multimolecular assemblies. The best characterized molecular chaperone is the bacterial ''chaperonin'', GroEL, and many of the details of the mechanism through which this molecule functions are now well understood (Hartl, 2000). Of particular significance for GroEL is the fact that it contains a cavity in which

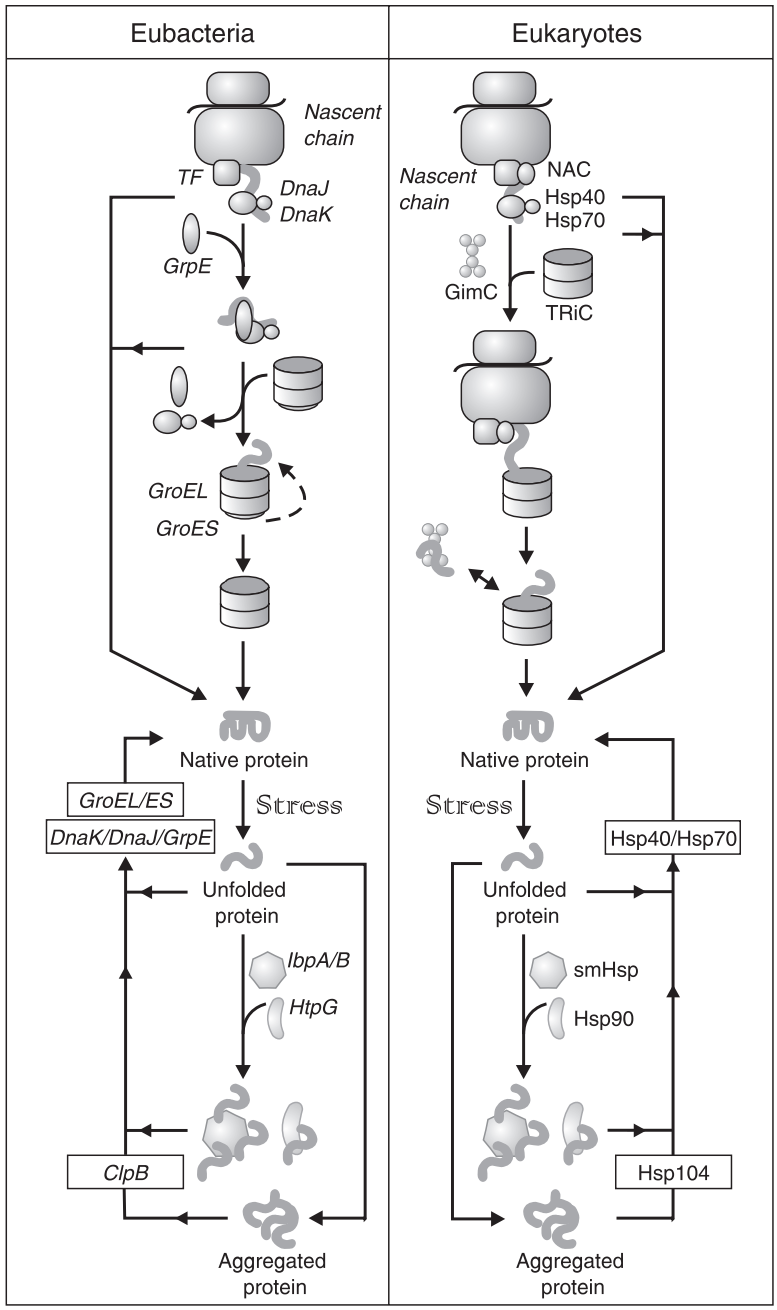

Figure 7.1 Schematic representation of the roles of a variety of molecular chaperones in the folding of proteins in bacterial and eukaryotic cells. As well as illustrating the sequential action of different types of molecular chaperone, this figure illustrates the range of different roles they play in avoiding misfolding and aggregation under normal and stress conditions. From (Hartl, 2000).

incompletely folded polypeptide chains can be sequestered and protected from the external environment. As well as molecular chaperones, there are several classes of folding catalyst that act to accelerate steps in the folding process which can otherwise be extremely slow (Gething and Sambrook, 1992). The most important are peptidylprolyl isomerases, which increase the rate of *cis/trans* isomerization of peptide bonds involving proline residues, and protein disulphide isomerases which enhance the rate of formation and reorganisation of disulphide bonds within proteins (Gething and Sambrook, 1992; Hartl and Hayer-Hartl, 2002). In addition there are quality control and degradation mechanisms, such as the unfolded protein response, to ensure that irretrievably misfolded or damaged proteins are eliminated before any harm can ensue (Sherman and Goldberg, 2001).

The nature of protein misfolding

The term ''protein folding'' is universally accepted to refer to the process that ultimately results in the generation of the native structure from a completely or partially unfolded state. In this article the word ''misfolding'' will be used to describe processes that result in a protein acquiring a sufficient number of persistent non-native interactions or characteristics such that its overall structure and properties are significantly affected. Thus, an intermediate populated during the normal productive folding process will not usually be described as misfolded even though it may have a number of non-native contacts between specific residues. If a protein were to find its way into a compact relatively stable structure with a non-native topology, however, such a species would normally be described as misfolded as its essential architecture, and almost invariably its physical and biological behavior, are demonstrably different from those of the native state. Therefore, if a protein assembles into aggregates such as amyloid fibrils (see below), this process will be described as misfolding as the specific interactions that determine such structures usually differ very significantly from those that stabilize the native state and the resulting folds of the polypeptide chains are, at least in many regions, quite distinct in the two forms (Dobson, 2002). Moreover, in this example, the properties of the species formed are very different from those of the functional state of the protein. Some assembly process will not, however, be described as misfolding on this definition, for example those that give rise to biologically functional species such as tubulin or indeed the process of protein crystallization provided that the native state is preserved within the crystals.

The fact that proteins can misfold in vitro in the manner defined above is very well established. For example, one of the major problems in generating recombinant protein for research purposes or for use in biotechnology is recovering significant quantities of active protein from refolding procedures carried out in the laboratory. The fact that misfolding, and its consequences such as aggregation, can be an important feature of the behavior of proteins in vivo is clearly evident from observations that the levels of certain proteins are increased very substantially during cellular stress (Figure 7.1). Indeed, many molecular chaperones were first recognized in prokaryotic (bacterial) systems which had been subjected to stress generated by exposure to elevated temperatures, as their nomenclature as HSPs (heat shock proteins) indicates (Pelham, 1968). As well as assisting folding and preventing incompletely folded chains from forming misfolded species, it is clear that some molecular chaperones are able to rescue proteins that have misfolded and enable them to have a second chance to fold correctly. There are also examples of molecular chaperones that are known to be able to solubilize misfolded aggregates under at least some circumstances (Parsell *et al.*, 1994). Such active intervention requires energy, and not surprisingly adenosine triphosphate (ATP) is required for many of the molecular chaperones to function correctly (Hartl and Hayer-Hartl, 2002). Despite the fact that many

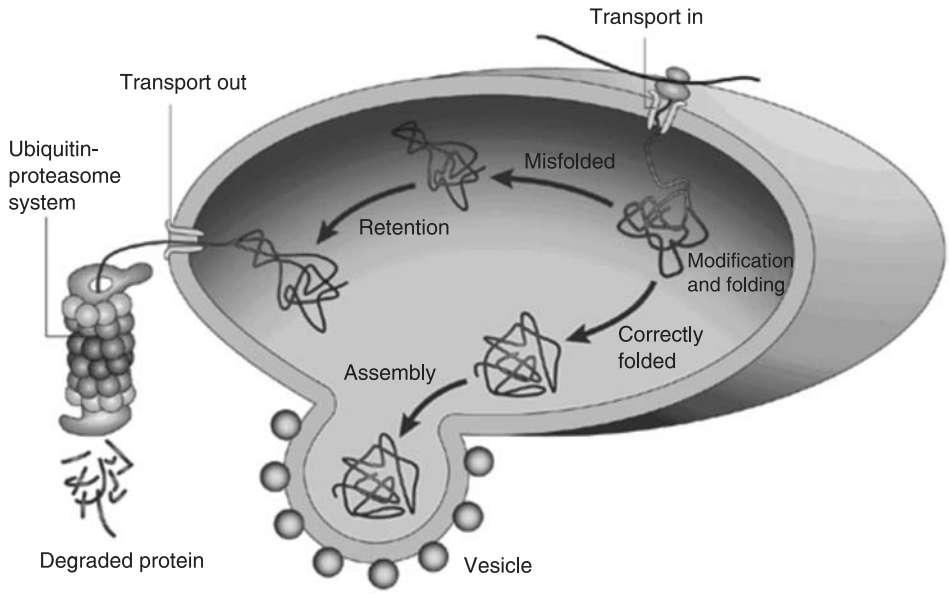

Figure 7.2 Folding in the endoplasmic reticulum. Newly synthesized proteins designed for export are transported into the endoplasmic reticulum (ER) where they undergo a series of modifications, such as glycosylation, and are helped to fold into their correct three-dimensional structures by molecular chaperones (not shown). Correctly folded proteins are then transported to the Golgi complex, encapsulated in membranes (called vesicles), and delivered to their final destination. Incorrectly folded proteins, however, are detected by a quality control mechanism, retained in the ER and sent along another pathway (the unfolded protein response) to be degraded by the ubiquitin-proteasome system. From (Dobson, 2003*a*), adapted from (Kaufman *et al.*, 2002).

molecular chaperones are expressed at high levels only in stressed systems, it is clear that they have a critical role to play in all organisms even when present at lower levels under normal physiological conditions.

In eukaryotic systems, many of the proteins which are synthesized in a cell are destined for secretion to the extracellular environment. These proteins are translocated into the ER where folding takes place prior to secretion through the Golgi apparatus. The ER contains a wide range of molecular chaperones and folding catalysts to promote efficient folding, and in addition the proteins involved are subject to a stringent quality control in the ER (Hammond and Helenius, 1995). The quality control mechanism can involve a complex series of glycosylation and deglycosylation processes and acts to prevent misfolded proteins from being secreted from the cell, (Figure 7.2). Once recognized, unfolded and misfolded proteins are targeted for degradation through the ubiquitin–proteosome pathway (Kaufman *et al.*, 2002). The details of how these regulatory systems operate provide remarkable evidence of the stringent mechanisms that biology has established to ensure that misfolding, and its consequences, are minimized. The importance of such a process is underlined by the fact that recent experiments indicate that up to half of all polypeptide chains fail to satisfy the quality control mechanism in the ER, and for some proteins the success rate is even lower (Schubert *et al.*, 2000). Like the ''heat shock response'' in the cytoplasm, the ''unfolded protein response'' in the ER is also upregulated during stress and, as we shall see below, is strongly linked to the avoidance of misfolding diseases.

Table 7.1. Representative protein folding diseases [From Dobson, 2001]

Disease	Protein	Site of folding
Hypercholesterolaemia	Low-density lipoprotein receptor	ER
Cystic fibrosis	Cystic fibrosis trans-membrane regulator	ER
Phenylketonuria	Phenylalanine hydroxylase	cytosol
Huntington's disease	Huntingtin	cytosol
Marfan syndrome	Fibrillin	ER
Osteogenesis imperfecta	Procollagen	ER
Sickle cell anaemia	Haemoglobin	cytosol
αl-antitrypsin deficiency	αl-antitrypsin	ER
Tay-Sachs disease	β-hexosaminidase	ER
Scurvy	Collagen	ER
Alzheimer's disease	Amyloid β-peptide/tau	ER
Parkinson's disease	α-synuclein	cytosol
Scrapie/Creutzfeldt-Jakob disease	Prion protein	ER
Familial amyloidoses	Transthyretin/lysozyme	ER
Retinitis pigmentosa	Rhodopsin	ER
Cataracts	Crystallins	cytosol
Cancer	p53	cytosol

Protein misfolding and its consequences for disease

Folding and unfolding are the ultimate ways of generating and abolishing specific cellular activities, and unfolding is also the precursor to the ready degradation of a protein (Matouschek, 2003). Moreover, it is increasingly apparent that some events in the cell, such as translocation across membranes, can require proteins to be in unfolded or partially folded states. Processes as apparently diverse as trafficking, secretion, the immune response and the regulation of the cell cycle are in fact now recognized to be directly dependent on folding and unfolding (Radford and Dobson, 1999). It is not surprising therefore that failure to fold correctly, or to remain correctly folded, will give rise to the malfunctioning of living systems and therefore to disease. Indeed, it is becoming increasingly evident that a wide range of human diseases is associated with aberrations in the folding process (Table 7.1) (Dobson, 2001; Thomas *et al.*, 1995). Some of these diseases (e.g. cystic fibrosis) can be attributed to the simple fact that if proteins do not fold correctly they will not able to exercise their proper function; such misfolded species are often degraded rapidly within the cell. Other (e.g. disorders associated with α1-antitrypsin) result from the failure of proteins to be trafficked to the appropriate organs in the body (Carrell and Lomas, 2002; Powell and Zeitlin, 2002). In other cases, misfolded proteins escape all the protective mechanisms discussed above and form intractable aggregates within cells or in extracellular space. An increasing number of pathologies, including Alzheimer's and Parkinson's diseases, the spongiform encephalopathies and late onset diabetes, are known to be directly associated with the deposition of such aggregates in tissue, Table 7.2 (Dobson, 2001; Koo *et al.*, 1999; Pepys, 1995; Selkoe, 2003; Thomas *et al.*, 1995). Diseases of this type are amongst the most debilitating, socially disruptive and costly in the modern world, and they are becoming increasingly prevalent as our societies age as new agricultural, dietary and medical practices are adopted (Dobson, 2002).

One of the most characteristic features of many of the aggregation diseases is that they give rise to

Table 7.2. Examples of diseases associated with amyloid deposition [Adapted from (Dobson, 2001)]

Clinical syndrome	Fibril component	a: organ limited b: systemic
Alzheimer's disease	Aβ-peptide, 1–42, 1–43	a
Spongiform encephalopathies	Full length prion or fragments	a
Primary systemic amyloidosis	Intact light chain or fragments	b
Secondary systemic amyloidosis	76-residue fragment of amyloid A protein	b
Familial amyloidotic polyneuropathy I	Transthyretin variants and fragments	b
Senile systemic amyloidosis	Wild-type transthyretin and fragments	b
Hereditary cerebral amyloid angiopathy	Fragment of cystatin-C	a
Haemodialysis-related amyloidosis	β2-microglobulin	b
Familial amyloidotic polyneuropathy II	Fragments of apolipoprotein A-I	b
Finnish hereditary amyloidosis	71-residue fragment of gelsolin	b
Type II diabetes	Fragment of islet-associated polypeptide	a
Medullary carcinoma of the thyroid	Fragments of calcitonin	a
Atrial amyloidosis	Atrial natriuretic factor	a
Lysozyme amyloidosis	Full length lysozyme variants	b
Insulin-related amyloidosis	Full length insulin	b
Fibrinogen α-chain amyloidosis	Fibrinogen α-chain variants	b

the deposition of proteins in the form of amyloid fibrils and plaques (Selkoe, 2003). Such deposits can form in the brain, in vital organs such as the liver and spleen, or in skeletal tissue, depending on the disease involved. In the case of neurodegenerative diseases, the quantity of such aggregates can be almost undetectable in some cases, whilst in systemic diseases kilograms of protein can be found in such deposits. Each amyloid disease involves primarily the aggregation of a specific peptide or protein although a range of other components including other proteins and carbohydrates is also incorporated into the deposits when they form in vivo. The characteristics of the soluble forms of the twenty or so proteins involved in the well-defined amyloidoses are very varied – they range from intact globular proteins to largely unstructured peptide molecules – but the aggregated forms have many common characteristics in common (Sunde and Blake, 1997). Thus, amyloid deposits all show specific optical properties (such as birefringence) on binding certain dye molecules notably Congo red; these properties have been used in diagnosis for over a century. The fibrillar structures that are characteristic of many of this type of aggregate have very similar morphologies (long, unbranched and often twisted structures a few nm in diameter) and a characteristic ''cross-beta'' X-ray fibre diffraction pattern (Sunde and Blake, 1997). The latter reveals that the organised core structure is composed of β-sheets having strands running perpendicular to the fibril axis, (Figure 7.3) (Jimenez *et al.*, 1999). Fibrils having the essential characteristics of those found in ex vivo deposits can be reproduced in vitro from the component proteins under appropriate conditions, showing that they can self-assemble without the need for other components.

For many years it was generally assumed that the ability to form amyloid fibrils with the characteristics described above was limited to a relatively small number of proteins, largely those seen in disease states, and that these proteins must possess specific sequence motifs encoding the amyloid core structure. Recent studies have suggested, however, that the ability of polypeptide chains to form such structures is common, and indeed can be considered a generic feature of polypeptide chains (Chiti *et al.*, 1999; Dobson, 1999*b*; 2003*b*). The most direct evidence for the latter statement is

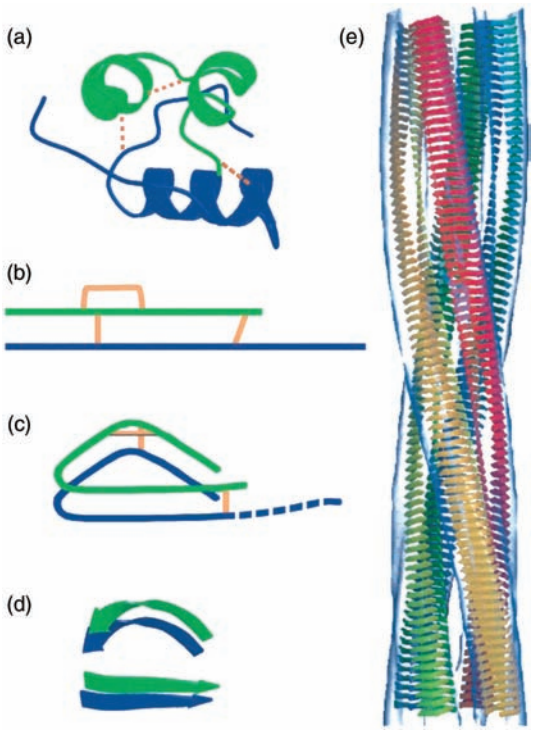

Figure 7.3 A molecular model of an amyloid fibril. This model is derived from cryo-EM analysis of fibrils grown from bovine insulin where native topology is indicated in (a). The two chains (A, green and B, blue) are connected by three disulphide bridges (b). A possible topology for insulin in the fibrils is illustrated in (c) whilst (d) indicates how strands could be assembled in a complete fibril, a model for what is shown in (e). The fibril consists of four ''protofilaments'' that twist around one another to form a regular pattern with a diameter of approximately 10 nm.

that fibrils can be formed under appropriate conditions by many different proteins that are not associated with disease, including such well-known proteins as myoglobin (Fandrich *et al.*, 2001), and also by homopolymers such as polythreonine or polylysine (Fandrich and Dobson, 2002). Remarkably, fibrils of similar appearance to those containing large proteins can be formed by peptides with just a handful of residues (Lopez De La Paz *et al.*, 2002). One can consider that amyloid fibrils are highly organised structures (effectively one-dimensional crystals) adopted by an unfolded polypeptide chain when it behaves as a typical polymer molecule; similar types of structure can, for example, be formed by synthetic polymers. The essential features of such structures are therefore determined by the physicochemical properties of the polymer chain. As with other highly organized materials (including crystals) whose structures are based on repetitive long-range interactions, the most stable structures are usually those consisting of a single type of molecular species (e.g. a specific peptide sequence) where such interactions can be optimized (Dobson, 2003*b*).

The structure and mechanism of amyloid formation

Studies of amyloid fibrils formed by both disease-associated and other peptides and proteins has enabled many of the features of these structures to be defined (Fandrich and Dobson, 2002; Lopez De La Paz *et al.*, 2002; Selkoe, 2003; Serpell *et al.*, 2000; Wille *et al.*, 2002), although no complete structure has yet been determined in atomic detail. It is clear that the core structure of the fibrils is stabilized primarily by interactions, particularly hydrogen bonds, involving the atoms of the extended polypeptide main chain. As the main chain is common to all polypeptides, this observation explains why fibrils formed from polypeptides of very different amino acid sequences are similar in appearance. The side chains are likely to be incorporated in whatever manner is most favourable for a given sequence within the amyloid structures; they appear to affect the details of the fibrillar assembly but not the fundamental structure of the fibril (Chamberlain *et al.*, 2000). In addition, the proportion of a polypeptide chain that is incorporated in the core structure can vary substantially; in some cases only a handful of residues may be involved in such structure with the remainder of the chain associated in some other

manner with the fibrillar assembly. This generic type of structure contrast strongly with the globular structures of most natural proteins that result from the uniquely favorable packing of a given set of side chains into a particular fold (Dobson, 1999*b*). On the conceptual model outlined here, structures such as amyloid fibrils occur because interactions associated with the highly specific packing of the side chains can sometimes override the conformational preferences of the main chain (Dobson, 1999*b*; 2003*b*; Fandrich *et al.*, 2001; Fandrich and Dobson, 2002). The strands and helices so familiar in the structures of native proteins are then the most stable structures that the main chain can adopt in the folds that are primarily defined by the side chain interactions. If the solution environment (pH, temperature, concentration etc.) in which the molecules are found is, however, such that these side chain interactions are insufficiently stable, the structures can unfold and may then, at least under some circumstances, reassemble in the form of amyloid fibrils.

In order to understand the occurrence and significance of amyloid formation in biological systems it is essential to establish the mechanism by which such structures are assembled from the soluble precursor species. In globular proteins the polypeptide main chain is largely buried within the folded structure. Conditions that favour formation of any form of aggregate by such species are therefore ones in which the molecules involved are at least partially unfolded, as occurs for example at low pH or elevated temperature (Kelly, 1998). Because of the importance of the globular fold in preventing aggregation, the fragmentation of proteins, through proteolysis or other means, is another ready mechanism to stimulate amyloid formation. Indeed, many amyloid disorders, including Alzheimer's disease, involve aggregation of fragments of larger precursor proteins that are unable to fold in the absence of the remainder of the protein structure, see Table 7.2. Experiments in vitro indicate that the formation of fibrils, by appropriately destabilized or fragmented proteins, is then generally characterized by a lag phase followed by a period of relatively rapid growth (Harper and Lansbury, 1997). Behavior of this type is typical of nucleated processes such as crystallization; as with crystallization, the lag phase in amyloid formation can generally be eliminated by addition of preformed fibrils to fresh solutions, a process known as seeding (Harper and Lansbury, 1997). Although the specific events taking place during fibril growth are not yet elucidated in any detail, it is becoming possible to simulate the overall kinetic profiles using relatively simple models that incorporate well-established principles of nucleated processes (Padrick and Miranker, 2002).

One of the key findings of studies of the formation of amyloid fibrils is that there are many common features in the behavior of the different systems that have been examined (Caughey and Lansbury, 2003; Harper and Lansbury, 1997; Koo *et al.*, 1999; Nettleton *et al.*, 2000). The first phase of the aggregation process appears generally to involve the formation of more or less disordered oligomeric species as a result of relatively non-specific interactions, although in some cases specific structural transitions, such as domain swapping (Schlunegger *et al.*, 1997), may be involved if such processes increase the rate of aggregation. The earliest species visible by electron or atomic force microscopy often resemble small bead-like structures, sometimes described as amorphous aggregates or as micelles. These early ''prefibrillar aggregates'' then appear to transform into species with more distinctive morphologies, sometimes described as ''protofibrils'' or ''protofilaments'' (Bitan *et al.*, 2003; Caughey and Lansbury, 2003; Jimenez *et al.*, 2002). These structures are commonly short, thin, sometimes curly, fibrillar species that are thought to be able to assemble into mature fibrils, perhaps by lateral association accompanied by some degree of structural reorganization (Bouchard *et al.*, 2000), Figure 7.4. The extent to which aggregates dissolve and reassemble into more regular structures is involved at the different stages of fibril assembly is not clear, but it could well be generally important under the slow growth conditions in which the

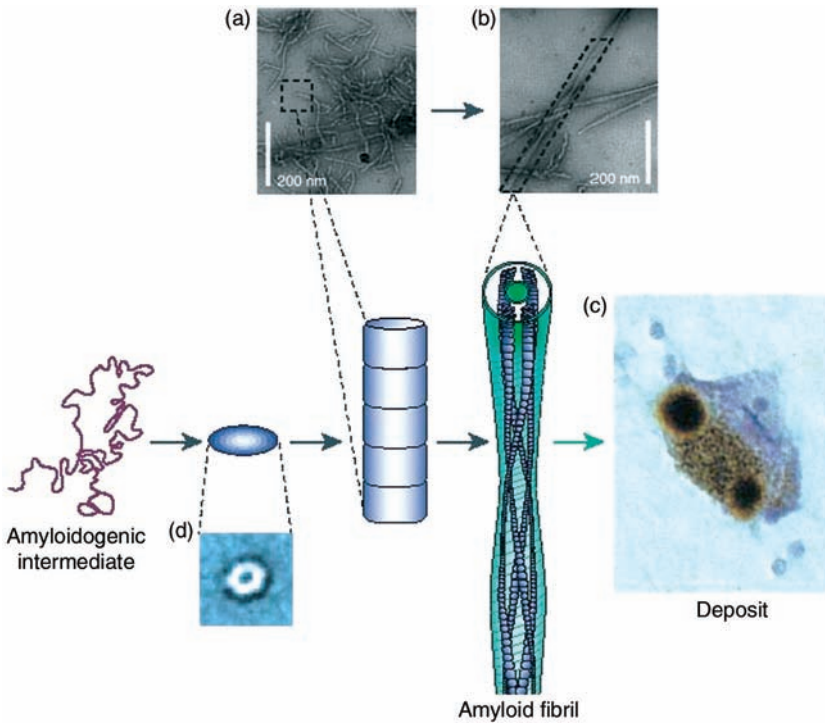

Figure 7.4 Schematic representation of the formation of amyloid fibrils and pathological deposits. An unfolded or partially unfolded peptide or protein initially forms small soluble aggregates that then assemble to form a variety of ''pre-fibrillar'' species, some or all of which appear to be toxic to living cells. These species undergo a series of additional assembly and reorganisational steps to give ordered amyloid fibrils of the type illustrated in Figure 7.3. Pathological deposits such as the Lewy bodies associated with Parkinson's disease, frequently contain large assemblies of fibrillar aggregates, along with other components including molecular chaperones (Spillantini *et al.*, 1997). From (Dobson, 2003*b*).

most highly structured fibrils are formed. The earliest aggregates are likely in most situations to be relatively disorganized structures which expose to the outside world a variety of segments of the protein that are normally buried in the globular state. In other cases, however, such species appear to adopt distinctive structures, including the well-defined ''doughnut'' shaped species that have been seen for a number of systems, (Figure 7.4) (Caughey and Lansbury, 2003; Lashuel *et al.*, 2002).

Although the ability of polypeptides to form amyloid fibrils appears generic, the propensity to do so varies very dramatically with the amino acid composition and sequence. At the most fundamental level, some types of amino acid are much more soluble than others such that the concentration that is required for aggregation to occur will be much greater for some polypeptides than for others. In addition the aggregation process, like crystallization, needs to be nucleated and the rates at which this process takes place can be highly dependent on many different factors. It is clear that even single changes of amino acids in protein sequences can change the rates at which the unfolded polypeptide chains aggregate by an order of magnitude or more (Chiti *et al.*, 2003). Moreover, it has recently proved possible to correlate changes in aggregation rates caused by mutations with changes in simple properties that result from such substitutions, such as charge,

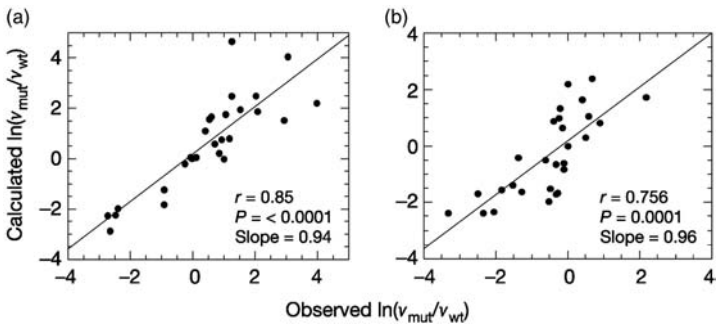

Figure 7.5 Rationalization of the effects of mutations on the aggregation rates of peptides and proteins. The experimental aggregation rates of a variety of short peptides or natively unfolded proteins, including amylin, amyloid β-peptide, tau and α-synuclein (see Tables 7.1 and 7.2 for a summary of the diseases with which these are associated) are shown plotted against rates calculated from an algorithm derived from extensive mutational studies of the protein acylphosphatase; the analagous plot for the data for acylphosphatase is shown in (b). The correlation shown in (a) argues strongly for a common mechanism for amyloid formation and provides a platform both to predict the effects of natural mutations and to design polypeptides with altered aggregation properties. From (Chiti *et al.*, 2003).

secondary structure propensities and hydrophobicity (Chiti *et al.*, 2003). As this correlation has been found to hold for a wide range of different sequences, (Figure 7.5), it strongly endorses the concept of a common mechanism of amyloid formation. In accord with such ideas, those proteins that are completely or partially unfolded under normal conditions in the cell have sequences which are likely to have very low propensities to aggregate. An interesting and potentially important additional observation is that specific regions of the polypeptide chain appear to be responsible for nucleating the aggregation process (Chiti *et al.*, 2002). Interestingly, the residues that nucleate the folding of a globular protein seem to be distinct from those that nucleate its aggregation into amyloid fibrils (Chiti *et al.*, 2002). Such a characteristic, which may reflect the different nature of the partially folded species that initiate the assembly processes, offers the opportunity for evolutionary pressure to select sequences that favor folding over aggregation.

A generic description of amyloid formation

It is clear that proteins can adopt different conformational states under different conditions and that there can be considerable similarities in this regard between different proteins. This situation can be summarized by using a schematic representation of the different states that are in principle accessible to a polypeptide chain, see Figure 7.6 (Dobson and Karplus, 1999; Dobson, 1999*a*; 2003*b*; Jimenez *et al.*, 2002; Schubert *et al.*, 2000). On such a model description, the state that a given system populates under specific conditions will depend on the relative thermodynamic stabilities of the different states (in the case of oligomers and aggregates the concentration will be a critical parameter) and the kinetics of the various interconversion processes. In this diagram, amyloid fibrils are included as just one of the types of aggregates that can be formed by proteins, although they have particular significance in that their highly organized hydrogen-bonded structure gives them unique kinetic stability. This type of representation emphasizes that complex biological systems have become robust by controlling and regulating the various states accessible to a given polypeptide chain at given times and under given conditions, just as they regulate and control the various chemical transformations that take place in the cell (Hartl, 2000). The latter is achieved primarily through enzymes, and the former by

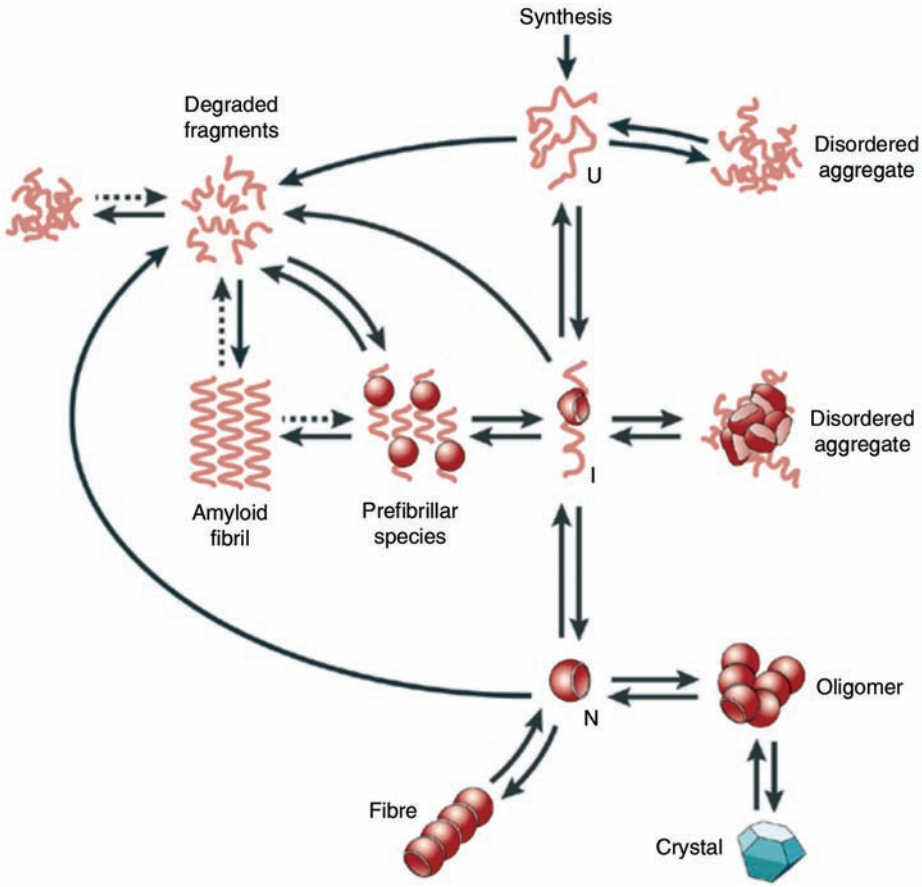

Figure 7.6 Schematic representation of some of the states accessible to a polypeptide chain following its synthesis on a ribosome. The relative populations of the different states depend on the kinetics and thermodynamics of the various equilibria shown in the diagram. In living systems the fate of a given molecule is closely regulated, using mechanisms such as those illustrated in Figures 7.1 and 7.2, rather as metabolic pathways in cells are controlled by enzymes and associated molecules such as cofactors. From (Dobson, 2003*a*).

means of the molecular chaperones and degradation mechanisms mentioned above. And just as the aberrant behavior of enzymes can cause metabolic disease, the aberrant behavior of the chaperone and other machinery regulating polypeptide conformations can contribute to misfolding and aggregation diseases (Macario and Conway de Macario, 2002).

The type of diagram shown in Figure 7.6 serves as a framework for understanding the fundamental molecular events that underlie the regulatory and quality control mechanisms, and the origins of the amyloid diseases which can result when these fail or are overwhelmed. In addition, it enables a rational approach to be taken to approaches for their prevention or treatment (Hartl, 2000). As we have discussed above, partially or completely unfolded polypeptides are highly aggregation-prone and represent the species that trigger amyloid formation. Such species are, however, inherent in the folding process and for this reason a variety of molecular chaperones is present in

abundance in the cellular compartments wherever such processes occur. It is possible, however, that chaperones also exist in environments where ab initio folding does not take place. Indeed, an extracellular chaperone (clusterin) has been identified recently (Wilson and Easterbrook-Smith, 2000). Nevertheless, it is undoubtedly important in the context of avoiding aggregation that proteins are correctly folded prior to their secretion from the cell, hence the need for a highly effective system of quality control in the ER. In this context it is interesting that the majority of the diseases associated with amyloid formation involve deposits that are extracellular in nature, although it is not yet clear exactly where their formation in fact takes place (Pepys, 1995). As biology is a dynamic process, there is a continuous need to degrade as well as synthesize proteins, and the degradatory mechanisms target misfolded as well as redundant proteins. It is during such processes, which require unfolding and proteolysis of polypeptide chains, that aggregation may be particularly likely. Degradation pathways, such as those of the ubiquitin-proteasome system, are therefore highly regulated in order to avoid the occurrence of such events (Bence *et al.*, 2001; Sherman and Goldberg, 2001).

In order to understand misfolding and aggregation diseases we need to know not just how such systems function efficiently, but also why they fail to do so under some circumstances (Bence *et al.*, 2001; Horwich, 2002; Stefani and Dobson, 2003). The effects of many pathogenic mutations can be particularly well understood from the schematic representation given in Figure 7.6. Many of the mutations associated with the familial deposition diseases increase the population of partially unfolded states either within or outside the cell by decreasing the stability or cooperativity of the native state (Booth *et al.*, 1997; Dumoulin *et al.*, 2003; Ramirez-Alvarado *et al.*, 2000). Cooperativity is in fact a crucial factor in enabling proteins to remain soluble as it ensures that even for a protein that is marginally stable, the equilibrium population of unfolded molecules or of unfolded regions of the polypeptide chain is minimal (Dobson, 1999*b*). Other familial diseases are associated with the accumulation of fragments of native proteins, produced by aberrant processing or incomplete degradation; such species are usually unable to fold into aggregation-resistant states. Other pathogenic mutations act by enhancing the propensities of such species to aggregate, for example by increasing their hydrophobicity or decreasing their charge (Chiti *et al.*, 2003). In the case of the transmissible encephalopathies, it is likely that ingestion of pre-aggregated states of an identical protein (e.g. by cannibalistic means or by contamination of surgical instruments) increases dramatically the rate of aggregation within the individual concerned, and hence underlies the mechanism of transmission (Chien *et al.*, 2003; Harper and Lansbury, 1997; Prusiner, 1997). Such seeding phenomena may also be at least part of the reason why some deposition conditions such as Alzheimer's disease progress so rapidly once the initial symptoms are evident (Harper and Lansbury, 1997; Perutz and Windle, 2001).

The fundamental origins of amyloid disease

Despite our increasing knowledge of the general principles that underlie protein misfolding diseases, the manner in which improperly folded proteins and aggregated proteins can generate pathological behavior is not yet understood in detail. In the case of systemic disease, the sheer mass of protein that can be deposited may physically disrupt the functioning of specific organs (Pepys, 1995). In other cases it may be that the loss of functional protein results in the failure of some crucial cellular process (Caughey and Lansbury, 2003). But for neurodegenerative disorders, such as Alzheimer's disease, it appears that the primary symptoms arise from the destruction of cells such as neurons by a ''toxic gain of function'' that results from the misfolding and aggregation process (Caughey and Lansbury, 2003; Koo *et al.*, 1999). It has recently become apparent

that the early prefibrillar aggregates of proteins associated with neurological disorders can be highly damaging to cells; by contrast the mature fibrils are relatively benign (Caughey and Lansbury, 2003; Walsh *et al.*, 2002). There is evidence, however, that the toxic nature of protein aggregates is not restricted to species formed from the peptides and proteins associated with pathological conditions. Experiments have recently indicated that prefibrillar aggregates of several proteins that are not connected with any known diseases have a comparable degree of toxicity to that of similar species formed from the Aβ-peptide (Bucciantini *et al.*, 2002). The concept that the effects of such aggregates on cells are more strongly related to their generic nature than to their specific sequence has recently been reinforced through experiments with polyclonal antibodies that cross-react with early aggregates of different peptides and proteins, and moreover inhibit their toxicity in cellular assays (Kayed *et al.*, 2003). It is possible that there are specific mechanisms for such toxicity, for example through the doughnut–shaped aggregates that resemble the toxins produces by bacteria which form pores in membranes and disrupt the ion balance in cells (Spillantini *et al.*, 1997). It is also possible that disorganized early aggregates, or even misfolded monomers, are toxic through a less specific mechanism, for example because exposed non-native hydrophobic surfaces may stimulate aberrant interactions with cell membranes or other cellular components (Oeppen and Vaupel, 2002; Stefani and Dobson, 2003; Svensson *et al.*, 1999).

Such findings raise the question as to how cellular systems are able to tolerate the intrinsic tendency of incompletely folded proteins to aggregate. The answer is almost certainly that under normal circumstances the molecular chaperones and other housekeeping mechanisms are remarkably efficient in ensuring that such potentially toxic species are neutralised (Hartl and Hayer-Hartl, 2002; Kayed *et al.*, 2003). Molecular chaperones of various types are able to shield hydrophobic regions, to unfold some forms of aggregates, or to alter the partitioning between different forms of aggregates. The latter mechanism, for example, could convert the precursors of amyloid fibrils into less intractable species allowing them to be refolded or disposed of by the cellular degradation systems. Indeed, evidence has been obtained that such a situation occurs with polyglutamine sequences associated with disorders such as Huntington's disease (Muchowski *et al.*, 2000). In this case the precursors of amyloid fibrils appear to be diverted into amorphous, and hence more readily degradable, species by the action of molecular chaperones. If such protective processes fail, it may be possible for potentially harmful species to be sequestered in relatively harmless forms such as inclusion bodies in bacteria or aggresomes in eukaryotic systems. Indeed, it has been suggested that the formation of mature amyloid fibrils, whose toxicity appears to be much lower than that of their precursors as we have discussed above, may itself represent a protective mechanism in some cases (Caughey and Lansbury, 2003; Koo *et al.*, 1999).

Most of the aggregation diseases do not, however, result from genetic mutations or infectious agents but are sporadic or associated with old age. The ideas summarized in this review offer a qualitative explanation of why such a situation arises. We have seen that all proteins have an inherent tendency to aggregate unless they are maintained in a highly regulated environment. Selective pressure during evolution has therefore resulted in protein molecules that are normally able to resist aggregation during our normal lifespan, enabling us to develop, pass on genes and to give appropriate protection to our offspring. Evolution, however, can only generate sequences and cellular environments that are adequate for those conditions under which selective pressure exists (Dobson, 2002). There is no reason to suppose that random mutations will in general reduce further the propensity to aggregate; indeed it will generally increase it, just as random mutations generally reduce the stability of native proteins. We can see, therefore, that our recent

ability to prolong life (Polverino de Laureto *et al.*, 2003) is leading to the proliferation of these diseases as the limitations in their ability to resist reversion to aggregates, including the intractable and damaging amyloid structures, become evident. It is intriguing, however, to speculate, that favorable mutations in aggregation-prone proteins might be the reason that some in the population do not readily succumb even in extreme old age to diseases such as Alzheimer's (Chiti *et al.*, 2003).

The link with aging is likely to involve more than just a greater probability of aggregation taking place as we get older. It is likely to be linked more fundamentally to the failure of the housekeeping mechanisms in our bodies under such circumstances (Csermely, 2001; Dobson, 2002; Macario and Conway de Macario, 2002). This failure may be in part be a result of the need for greater protective capacity in old age as aggregation becomes more prevalent, perhaps as a result of the increasing accumulation of misfolded and damaged proteins, leading to chaperone overload (Csermely, 2001). But as we age it is likely that the activity of our chaperone response and degradatory mechanisms declines, and that this results in the increasing probability that the protective mechanisms are overwhelmed. One reason for this decline, although there may be many, is that cells become less efficient in producing the ATP that is needed for the effective functioning of many chaperones. Similarly, the rapidity with which we have introduced practices that have not been experienced previously in history – including new agricultural practices (associated with the emergence of bovine spongiform encephalopathy (BSE) and variant Creutzfeldt-Jakob disease (CJD) (Ramirez-Alvarado *et al.*, 2000), a changing diet and lifestyle (associated with the prevalence of type 2 diabetes (Hoppener *et al.*, 2002) and new medical procedures (e.g. resulting in iatrogenic diseases such as CJD (Prusiner, 1997) and amyloid deposition in hemodialysis during which the concentration of β-2 microglobulin in serum increases (Gejyo *et al.*, 1986) – means that we have not had time to evolve effective protective mechanisms (Dobson, 2002). Thus we are now increasingly observing the limitations of our proteins and their environments to resist aggregation under conditions that differ from those under which natural selection has taken place.

Approaches to therapeutic intervention in amyloid disease

The progress made in understanding the underlying causes of the amyloid diseases is also leading to new approaches to their prevention and treatment. For proteins whose functional state is a tightly packed globular fold, we have seen that an essential first step in fibril formation is undoubtedly the partial or complete unfolding of the native structure that otherwise protects the aggregation-prone polypeptide backbone. Thus many of the familial forms of amyloid disease are associated with genetic mutations which decrease protein stability and so promote unfolding. In such cases, one approach to therapy is to find ways of stabilizing the native states of amyloidogenic protein variants. In accord with this proposal, a recent study has reported a series of small molecule analogues of the hormone thyroxine, the natural ligand of transthyretin, that act in this manner. Transthyretin is the protein associated with the most common type of systemic amyloidosis, and these potential drugs have been found to inhibit dramatically the rate at which the disease-associated variants aggregate at least in vitro (Hammarstrom *et al.*, 2003). In a similar approach, specific antibodies raised against lysozyme have been found to prevent amyloid fibril formation by pathogenic forms of this well-known antibacterial protein whose aggregation is linked to another systemic amyloidosis (Dumoulin *et al.*, 2003). Moreover, quinacrine, a drug previously used to combat malaria, has been found to limit dramatically the replication in cell cultures of the pathogenic form of the prion protein associated with CJD (May *et al.*, 2003). The finding that this

molecule interacts with the soluble form of the protein (Vogtherr *et al.*, 2003) suggests that a similar mechanism could be operating here. Quinacrine is already entering clinical trials, and determined efforts are now underway to find more potent forms of this compound (May *et al.*, 2003).

For those types of amyloid disease that result not from the aggregation of intact proteins but of proteolytic fragments, resulting from natural processing or from incomplete degradation, that are unable to fold in the absence of the remainder of the protein. A possible therapeutic strategy for such diseases would be to reduce the levels of the aggregation-prone peptides by inhibiting the enzymes through whose action the fragments are generated. An extremely important example of this approach involves the 40 or 42 residue Aβ-peptide derived from the amyloid precursor protein (APP) as the aggregation of this species is linked to Alzheimer's disease. Much is now known about the complex secretase enzymes whose role in processing APP gives rise to the Aβ-peptide (Wolfe, 2002), and a number of potent inhibitors have already been developed as potential therapeutics for this debilitating neurodegenerative disease (Vassar, 2002). Another way of reducing the levels of aggregation-prone species in an organism is to enhance their clearance from the body. A number of strategies based on this general idea have emerged recently, some of the most exciting of which involve the action of antibodies raised against specific amyloid-forming polypeptides. Of particular interest is the finding that active immunization with Aβ-peptides can result in the extensive clearance of amyloid deposits in transgenic mice that overexpress the Aβ-peptide. Although the first set of clinical trials with human Alzheimer victims was terminated as a result of an inflammatory response in a significant number of patients, the potential of this type of approach has been clearly demonstrated (Nicoll *et al.*, 2003). One possible variation on the original procedure is to use passive immunotherapy, in which antibodies are infused into, rather than generated within, the patient. There are in addition means of stimulating clearance other than by the direct action of antibodies. In a study of the effects of a ligand designed to bind to serum amyloid P (SAP), a protein that is associated with amyloid deposits in vivo and is known to inhibit natural clearance mechanisms, the levels of SAP in serum were reduced dramatically (Pepys *et al.*, 2002). It remains to be established how such molecules fare in ongoing clinical trials, but the general principle of using small molecules to target specific proteins for degradation is an extremely interesting development.

Enhancing the clearance of the amyloidogenic species may well be of particular importance in other forms of amyloid disorders that involve the deposition of peptides and proteins which, although intact, appear not to fold to globular structures even under normal physiological conditions. Examples of such cases include α-synuclein, whose aggregation is associated with Parkinson's disease, and amylin, involved in type 2 diabetes (Selkoe, 2003). Another strategy that has been proposed is to intervene in the aggregation process directly by means of small peptides or peptide analogues designed to bind tightly to fibrils as they form, hence blocking their further growth. A variety of such species is being explored, particularly in the context of Alzheimer's disease (Schenk, 2002). A potential problem with this approach is that inhibition of the formation of amyloid fibrils might occur at the expense of the small aggregates which are their precursors. Such an effect could be highly counterproductive as the latter appear to be the primary pathogenic agents, at least in the neurodegenerative forms of amyloid disease (Selkoe, 2003). However, provided that problems of this type can be avoided, this general approach could potentially be rather generally applicable.

It is a particularly satisfying aspect of research directed at understanding the origins of the amyloid disorders to find that so many different therapeutic approaches can be rationalized using the relatively simple framework illustrated in

Figure 7.6 (Dobson, 2004). Particularly as there are a number of different stages in the aggregation process where intervention is potentially possible, one can be optimistic that novel and efficaceous forms of treatment will emerge in the not too distant future. Moreover, with the increasing evidence for the generic nature of the amyloid conditions, it is interesting to speculate that there might be generic approaches to inhibiting a variety, perhaps even all, of the amyloid diseases with a single drug (Dobson, 2003*b*; Stefani and Dobson, 2003). That such an idea could, in principle, be realized is suggested by the remarkable report that antibodies raised against small aggregates of the Aβ-peptide not only recognize, but also suppress the toxicity of, similar aggregates formed from other proteins (Kayed *et al.*, 2003). An approach of this type would be particularly dramatic in the light of the number of different types of amyloid disorders which increasingly place the aging populations of the world at risk. Alternatively, one might even speculate that a means could be found to extend the period of time that our natural defences are able to combat aggregation, and hence defer the onset of these disorders. Looking perhaps even further into the future, our ability to design peptides and proteins whose tendencies to aggregate are reduced, often substantially, provides the basis for a range of exciting approaches involving gene therapy and stem cell techniques (Chiti *et al.*, 2003).

Concluding remarks

This article has discussed the state of our developing knowledge of the origin of amyloid disease from the perspective of our increasing understanding of the fundamental properties of proteins. We have stressed the fact that the structure and biological effects of amyloid deposition can be considered to be generic and that this fact can be rationalized at least in outline through the ideas of polymer science. The propensities of different sequences to aggregate under specific conditions, however, can vary very significantly and it is becoming possible to explain the origins of pathogenic behavior in terms of the factors that affect such propensities. The amyloid diseases can be though of as fundamentally resulting from the reversion of normally soluble proteins to the generic amyloid structure as the result of the failures of the mechanisms that have evolved to prevent such behavior. A range of novel approaches is, however, now being developed to combat the various amyloid diseases, and can to a large extent be rationalized in terms of intervention at different steps in the process of aggregation as represented in Figure 7.6. Many of these approaches show considerable promise for the treatment of specific diseases, and the generic nature of amyloid formation raises the exciting possibility that there could be generic approaches to preventing some or indeed all of the various members of the family of diseases. We can therefore be optimistic that the better understanding of protein misfolding and aggregation that is developing from recent research of the type discussed in this article, will enable us to devise increasingly effective strategies for drug discovery, based on rational arguments.

Acknowledgements

The ideas in this article have emerged from extensive discussions with outstanding students, research fellows and colleagues over many years. I am most grateful to all of them; they are too numerous to mention here but the names of many appear in the list of references. The content of this article is derived in part from a review article published in *Seminars in Cell and Developmental Biology* (2004) and a Perspective article published in *Science* (2004). The research of CMD is primarily supported by Programme Grants from the Wellcome Trust and the Leverhulme Trust, as well as by the BBSRC, EPSRC and MRC.

REFERENCES

Bence, N.F., Sampat, R.M. and Kopito, R.R. (2001). Impairment of the ubiquitin-proteasome system by protein aggregation. *Science*, **292**, 1552–5.

Bitan, G., Kirkitadze, M.D., Lomakin, A. *et al.* (2003). Amyloid beta-protein (Abeta) assembly: Abeta 40 and Abeta 42 oligomerize through distinct pathways. *Proc Natl Acad Sci USA*, **100**, 330–5.

Booth, D.R., Sunde, M., Bellotti, V. *et al.* (1997). Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. *Nature*, **385**, 787–93.

Bouchard, M., Zurdo, J., Nettleton, E.J., Dobson, C.M. and Robinson, C.V. (2000). Formation of insulin amyloid fibrils followed by FTIR simultaneously with CD and electron microscopy. *Protein Sci*, **9**, 1960–7.

Bucciantini, M., Giannoni, E., Chiti, F. *et al.* (2002). Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature*, **416**, 507–11.

Capaldi, A.P., Kleanthous, C. and Radford, S.E. (2002). Im7 folding mechanism: misfolding on a path to the native state. *Nat Struct Biol*, **9**, 209–16.

Carrell, R.W. and Lomas, D.A. (2002). Alpha1-antitrypsin deficiency–a model for conformational diseases. *N Engl J Med*, **346**, 45–53.

Caughey, B. and Lansbury, P.T. (2003). Protofibrils, pores, fibrils, and neurodegeneration: separating the responsible protein aggregates from the innocent bystanders. *Annu Rev Neurosci*, **26**, 267–98.

Chamberlain, A.K., MacPhee, C.E., Zurdo, J. *et al.* (2000). Ultrastructural organization of amyloid fibrils by atomic force microscopy. *Biophys J*, **79**, 3282–93.

Chien, P., DePace, A.H., Collins, S.R. and Weissman, J.S. (2003). Generation of prion transmission barriers by mutational control of amyloid conformations. *Nature*, **424**, 948–51.

Chiti, F., Stefani, M., Taddei, N., Ramponi, G. and Dobson, C.M. (2003). Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature*, **424**, 805–8.

Chiti, F., Taddei, N., Baroni, F. *et al.* (2002). Kinetic partitioning of protein folding and aggregation. *Nat Struct Biol*, **9**, 137–43.

Chiti, F., Webster, P., Taddei, N. *et al.* (1999). Designing conditions for in vitro formation of amyloid protofilaments and fibrils. *Proc Natl Acad Sci USA*, **96**, 3590–4.

Csermely, P. (2001). Chaperone overload is a possible contributor to 'civilization diseases'. *Trends Genet*, **17**, 701–4.

Dobson, C.M. (1999*a*). How do we explore the energy landscape for folding? In Fraunfelder, H., Deisenhofer, J. and Wolynes, P.G. ed., *Simplicity and Complexity in Proteins and Nucleic Acids*, pp. 15–37. Berlin: Dahlem University Press.

Dobson, C.M. (1999*b*). Protein misfolding, evolution and disease. *Trends Biochem Sci*, **24**, 329–32.

Dobson, C.M. (2001). The structural basis of protein folding and its links with human disease. *Philos Trans R Soc Lond B Biol Sci*, **356**, 133–45.

Dobson, C.M. (2002). Getting out of shape. *Nature*, **418**, 729–30.

Dobson, C.M. (2003*a*). Protein folding and disease: a view from the first Horizon Symposium. *Nat Rev Drug Discov*, **2**, 154–60.

Dobson, C.M. (2003*b*). Protein folding and misfolding. *Nature*, **426**, 884–90.

Dobson, C.M. (2004). Protein chemistry. In the footsteps of alchemists. *Science*, **304**, 1259–62.

Dobson, C.M. and Karplus, M. (1999). The fundamentals of protein folding: bringing together theory and experiment. *Curr Opin Struct Biol*, **9**, 92–101.

Dumoulin, M., Last, A.M., Desmyter, A. *et al.* (2003). A camelid antibody fragment inhibits the formation of amyloid fibrils by human lysozyme. *Nature*, **424**, 783–8.

Fandrich, M. and Dobson, C.M. (2002). The behaviour of polyamino acids reveals an inverse side chain effect in amyloid structure formation. *Embo J*, **21**, 5682–90.

Fandrich, M., Fletcher, M.A. and Dobson, C.M. (2001). Amyloid fibrils from muscle myoglobin. *Nature*, **410**, 165–6.

Gejyo, F., Homma, N., Suzuki, Y. and Arakawa, M. (1986). Serum levels of beta 2-microglobulin as a new form of amyloid protein in patients undergoing long-term hemodialysis. *N Engl J Med*, **314**, 585–6.

Gething, M.J. and Sambrook, J. (1992). Protein folding in the cell. *Nature*, **355**, 33–45.

Hammarstrom, P., Wiseman, R.L., Powers, E.T. and Kelly, J.W. (2003). Prevention of transthyretin amyloid disease by changing protein misfolding energetics. *Science*, **299**, 713–16.

Hammond, C. and Helenius, A. (1995). Quality control in the secretory pathway. *Curr Opin Cell Biol*, **7**, 523–9.

Hardesty, B. and Kramer, G. (2001). Folding of a nascent peptide on the ribosome. *Prog Nucleic Acid Res Mol Biol*, **66**, 41–66.

Harper, J. D. and Lansbury, P. T., Jr. (1997). Models of amyloid seeding in Alzheimer's disease and scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. *Annu Rev Biochem*, **66**, 385–407.

Hartl, F. U. and Hayer-Hartl, M. (2002). Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science*, **295**, 1852–8.

Hartl, U. (2000). In Pain, R. H. (ed.), *Mechanisms of protein folding*, Oxford: Oxford University Press.

Hoppener, J. W., Nieuwenhuis, M. G., Vroom, T. M., Ahren, B. and Lips, C. J. (2002). Role of islet amyloid in type 2 diabetes mellitus: consequence or cause? *Mol Cell Endocrinol*, **197**, 205–12.

Hore, P. J., Winder, S. L., Roberts, C. H. and Dobson, C. M. (1997). Stopped-flow photo-CIDNP observation of protein folding. *J Am Chem Soc*, **119**, 5049–50.

Horwich, A. (2002). Protein aggregation in disease: a role for folding intermediates forming specific multimeric interactions. *J Clin Invest*, **110**, 1221–32.

Jimenez, J. L., Guijarro, J. I., Orlova, E. *et al.* (1999). Cryo-electron microscopy structure of an SH3 amyloid fibril and model of the molecular packing. *Embo J*, **18**, 815–21.

Jimenez, J. L., Nettleton, E. J., Bouchard, M. *et al.* (2002). The protofilament structure of insulin amyloid fibrils. *Proc Natl Acad Sci USA*, **99**, 9196–201.

Kaufman, R. J., Scheuner, D., Schroder, M. *et al.* (2002). The unfolded protein response in nutrient sensing and differentiation. *Nat Rev Mol Cell Biol*, **3**, 411–21.

Kayed, R., Head, E., Thompson, J. L. *et al.* (2003). Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science*, **300**, 486–9.

Kelly, J. W. (1998). The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Curr Opin Struct Biol*, **8**, 101–6.

Koo, E. H., Lansbury, P. T., Jr. and Kelly, J. W. (1999). Amyloid diseases: abnormal protein aggregation in neurodegeneration. *Proc Natl Acad Sci USA*, **96**, 9989–90.

Lashuel, H. A., Hartley, D., Petre, B. M., Walz, T. and Lansbury, P. T., Jr. (2002). Neurodegenerative disease: amyloid pores from pathogenic mutations. *Nature*, **418**, 291.

Lopez De La Paz, M., Goldie, K., Zurdo, J. *et al.* (2002). De novo designed peptide-based amyloid fibrils. *Proc Natl Acad Sci USA*, **99**, 16052–7.

Macario, A. J. and Conway de Macario, E. (2002). Sick chaperones and ageing: a perspective. *Ageing Res Rev*, **1**, 295–311.

Matouschek, A. (2003). Protein unfolding–an important process in vivo? *Curr Opin Struct Biol*, **13**, 98–109.

May, B. C., Fafarman, A. T., Hong, S. B. *et al.* (2003). Potent inhibition of scrapie prion replication in cultured cells by bis-acridines. *Proc Natl Acad Sci USA*, **100**, 3416–21.

Muchowski, P. J., Schaffar, G., Sittler, A. *et al.* (2000). Hsp70 and hsp40 chaperones can inhibit self-assembly of polyglutamine proteins into amyloid-like fibrils. *Proc Natl Acad Sci USA*, **97**, 7841–6.

Nettleton, E. J., Tito, P., Sunde, M. *et al.* (2000). Characterization of the oligomeric states of insulin in self-assembly and amyloid fibril formation by mass spectrometry. *Biophys J*, **79**, 1053–65.

Nicoll, J. A., Wilkinson, D., Holmes, C. *et al.* (2003). Neuropathology of human Alzheimer disease after immunization with amyloid-beta peptide: a case report. *Nat Med*, **9**, 448–52.

Oeppen, J. and Vaupel, J. W. (2002). Demography. Broken limits to life expectancy. *Science*, **296**, 1029–31.

Padrick, S. B. and Miranker, A. D. (2002). Islet amyloid: phase partitioning and secondary nucleation are central to the mechanism of fibrillogenesis. *Biochemistry*, **41**, 4694–703.

Parsell, D. A., Kowal, A. S., Singer, M. A. and Lindquist, S. (1994). Protein disaggregation mediated by heat-shock protein Hsp104. *Nature*, **372**, 475–8.

Pelham, H. R. (1968). Speculations on the functions of the major heat shock and glucose-regulated proteins. *Cell*, **46**, 959–61.

Pepys, M. B. (1995). In Weatherall, D. J., Ledingham, J. G. and Warrel, D. A. (eds.), *Oxford textbook of medicine*, pp. 1512–24. Oxford: Oxford University Press.

Pepys, M. B., Herbert, J., Hutchinson, W. L. *et al.* (2002). Targeted pharmacological depletion of serum amyloid P component for treatment of human amyloidosis. *Nature*, **417**, 254–9.

Perutz, M. F. and Windle, A. H. (2001). Cause of neural death in neurodegenerative diseases attributable to expansion of glutamine repeats. *Nature*, **412**, 143–4.

Polverino de Laureto, P., Taddei, N., Frare, E. *et al.* (2003). Protein aggregation and amyloid fibril formation by

an SH3 domain probed by limited proteolysis. *J Mol Biol*, **334**, 129–41.

Powell, K. and Zeitlin, P.L. (2002). Therapeutic approaches to repair defects in deltaF508 CFTR folding and cellular targeting. *Adv Drug Deliv Rev*, **54**, 1395–408.

Prusiner, S.B. (1997). Prion diseases and the BSE crisis. *Science*, **278**, 245–51.

Radford, S.E. and Dobson, C.M. (1999). From computer simulations to human disease: emerging themes in protein folding. *Cell*, **97**, 291–8.

Ramirez-Alvarado, M., Merkel, J.S. and Regan, L. (2000). A systematic exploration of the influence of the protein stability on amyloid fibril formation in vitro. *Proc Natl Acad Sci USA*, **97**, 8979–84.

Schenk, D. (2002). Amyloid-beta immunotherapy for Alzheimer's disease: the end of the beginning. *Nat Rev Neurosci*, **3**, 824–8.

Schlunegger, M.P., Bennett, M.J. and Eisenberg, D. (1997). Oligomer formation by 3D domain swapping: a model for protein assembly and misassembly. *Adv Protein Chem*, **50**, 61–122.

Schubert, U., Anton, L.C., Gibbs, J. *et al.* (2000). Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature*, **404**, 770–4.

Selkoe, D.J. (2003). Folding proteins in fatal ways. *Nature*, **426**, 900–4.

Serpell, L.C., Blake, C.C. and Fraser, P.E. (2000). Molecular structure of a fibrillar Alzheimer's A beta fragment. *Biochemistry*, **39**, 13269–75.

Sherman, M.Y. and Goldberg, A.L. (2001). Cellular defenses against unfolded proteins: a cell biologist thinks about neurodegenerative diseases. *Neuron*, **29**, 15–32.

Spillantini, M.G., Schmidt, M.L., Lee, V.M. *et al.* (1997). Alpha-synuclein in Lewy bodies. *Nature*, **388**, 839–40.

Stefani, M. and Dobson, C.M. (2003). Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *J Mol Med*, **81**, 678–99.

Sunde, M. and Blake, C. (1997). The structure of amyloid fibrils by electron microscopy and X-ray diffraction. *Adv Protein Chem*, **50**, 123–59.

Svensson, M., Sabharwal, H., Hakansson, A. *et al.* (1999). Molecular characterization of alpha-lactalbumin folding variants that induce apoptosis in tumor cells. *J Biol Chem*, **274**, 6388–96.

Thomas, P.J., Qu, B.H. and Pedersen, P.L. (1995). Defective protein folding as a basis of human disease. *Trends Biochem Sci*, **20**, 456–9.

Vassar, R. (2002). Beta-secretase (BACE) as a drug target for Alzheimer's disease. *Adv Drug Deliv Rev*, **54**, 1589–602.

Vendruscolo, M., Zurdo, J., MacPhee, C.E. and Dobson, C.M. (2003). Protein folding and misfolding: a paradigm of self-assembly and regulation in complex biological systems. *Philos Transact A Math Phys Eng Sci*, **361**, 1205–22.

Vogtherr, M., Grimme, S., Elshorst, B. *et al.* (2003). Antimalarial drug quinacrine binds to C-terminal helix of cellular prion protein. *J Med Chem*, **46**, 3563–4.

Walsh, D.M., Klyubin, I., Fadeeva, J.V. *et al.* (2002). Naturally secreted oligomers of amyloid beta protein potently inhibit hippocampal long-term potentiation in vivo. *Nature*, **416**, 535–9.

Wille, H., Michelitsch, M.D., Guenebaut, V. *et al.* (2002). Structural studies of the scrapie prion protein by electron crystallography. *Proc Natl Acad Sci USA*, **99**, 3563–8.

Wilson, M.R. and Easterbrook-Smith, S.B. (2000). Clusterin is a secreted mammalian chaperone. *Trends Biochem Sci*, **25**, 95–8.

Wolfe, M.S. (2002). Secretase as a target for Alzheimer's disease. *Curr Top Med Chem*, **2**, 371–83.

8

Aging and disease

Thomas T. Perls

Is achieving extreme old age worthwhile? the centenarian phenotype

Average life expectancy has markedly increased over the past century. In 1900 average life expectancy was 46 years and in the United States it is currently almost 79 years. The age 85+ group is the fastest growing segment of our population and within that group, the number of centenarians is growing even faster. Whether mortality declines have been accompanied by health improvements among the elderly has been a matter of debate. Some authors have suggested that mortality declines have led to increased prevalence of frailty among older survivors because treatment of existing diseases simply postpones death to older ages (Gruenberg, 1977; Kramer, 1983; Olshansky and Ault, 1986), while others suggest that mortality declines have led to a compression of morbidity (Fries, 1980). Early US evidence from the 1970s was generally consistent with the idea that health among the elderly had deteriorated (Crimmins and Ingegneri, 1993; Verbrugge, 1984), while more recent evidence provides a somewhat more optimistic view (Freedman and Soldo, 1994; Manton *et al.*, 1997). Changes in disability prevalence over time, however, have varied by type and severity and by the data source used for trend analyses (Waidmann and Manton, 1998).

Centenarians studied by Hitt *et al.* appear to fit the pattern of postponement of disability to very advanced ages. In a population-based sample, it was observed that 90% of centenarians were independently functioning at age 92 (Hitt *et al.*, 1999). Most subjects experienced a decline in their cognitive function only in the last three to five years of their lives (Perls, 1997; Silver *et al.*, 1998). Upon further examination of the ages of onset for ten common age-associated diseases (hypertension, heart disease, diabetes, stroke, non-skin cancer, skin cancer, osteoporosis, thyroid condition, Parkinson's disease, chronic obstructive pulmonary disease and cataracts) among 424 centenarians (323 males and 101 females), the subjects were noted to fit into three morbidity profiles: ''survivors,'' ''delayers'' and ''escapers'' (Evert *et al.*, 2003). Survivors, individuals who were diagnosed with age-related illness prior to age 80, accounted for 24% of the male and 43% of the female centenarians ($p = 0.0009$). Delayers, individuals who delayed the onset of age-related diseases until at least age 80, accounted for 44% of the male and 42% of female centenarians. Escapers, individuals who attained the hundredth year of life without the diagnosis of an age-related disease, accounted for 32% of the male and 15% of the female centenarians ($p = 0.0003$). In this study, the prevalence of coronary heart disease was 46% for men with a median age of diagnosis of 92 years and 39% for women with a median age of diagnosis of 93 years (Evert *et al.*, 2003). The finding of a substantial number of centenarians fitting a ''survivor'' profile may be inconsistent with the compression of morbidity hypothesis. That most centenarians appear to be functionally independent through their early nineties suggests the

possibility that ''survivors'' and ''delayers'' are better able to cope with illnesses and remain functionally independent. Thus, in the case of centenarians, it may be more accurate to note a compression of disability rather than morbidity. This is not the case, as would be expected, with illnesses associated with high mortality risks. When examining only the most lethal diseases of the elderly, including heart disease, non-skin cancer, and stroke, 87% of males and 83% of females delayed or escaped these diseases.

These results suggest that there may be multiple routes to achieving exceptional longevity and that there are gender differences according to which route is taken. These routes represent different phenotypes and thus likely different genotypes of centenarians. For example, differential coping abilities could be due to varying degrees of genetic homeostasis. If so, then what are the genetic (homeostatic) mechanisms involved in this (e.g. heterozygosity or dominance?). The identification of three subtypes of centenarians, Survivors, Delayers, and Escapers, provides direction for future study into factors that determine exceptional longevity (Evert *et al.*, 2003).

The relative contributions of genes, environment and luck to how we age

The relative contribution of environmental and genetic influences to life expectancy has been a source of debate. Assessing heritability in 10 505 Swedish twin pairs reared together and apart, Ljungquist and colleagues attributed 35% of the variance in longevity to genetic influences and 65% of the variance to non-shared environmental effects (Ljungquist *et al.*, 1998). Other twin studies indicate heritability estimates of life expectancy between 25% and 30% (Herskind *et al.*, 1996; McGue *et al.*, 1993). A study of 1655 old order Amish subjects born between 1749 and 1890 and surviving beyond age 30 years resulted in a heritability calculation for life span of 0.25 (Mitchell *et al.*, 2001). These studies support the contention that the life spans of average humans with their average set of genetic polymorphisms are differentiated primarily by their habits and environments. Supporting this idea is a study of Seventh Day Adventists. In contrast to the American average life expectancy of 78 years, the average life expectancy of Seventh Day Adventists is 88 years. Because of their religious beliefs, members of this religious faith maintain optimal health habits, such as not smoking, a vegetarian diet, regular exercise and maintenance of a lean body mass, that translate into the addition of 10 years to their average life expectancy compared with other Americans (Fraser and Shavlik, 2001). Given that in the United States, 75% of persons are overweight and one third are obese (Fontaine *et al.*, 2003) far too many persons still use tobacco (Wechsler *et al.*, 1998), and far too few persons regularly exercise (Wei *et al.*, 1999), it is no wonder that our average life expectancy is about 10 years less than what our average set of genes should be able to achieve for us.

Of course there are exceptions to the rule. There are individuals who have genetic profiles with or without prerequisite environmental exposures that predispose them to diseases at younger ages. There is also a component of luck which, good or bad, plays a role in life expectancy. Finally, there is the possibility that there exist genetic and environmental factors which facilitate the ability to live to ages significantly older than those which the average set of genetic and environmental exposures would normally allow. Because the oldest individuals in the twin studies were in their early to mid eighties, those studies provide information about heritability of average life expectancy, but not of substantially older ages, for example age 100 years and older. As discussed below, to survive the 15 or more years beyond what our average set of genetic variations is capable of achieving for us, it appears that people need to have benefited from a relatively rare combination of what might be not so rare environmental, behavioral and genetic characteristics and that a number of these factors appear to be shared within families.

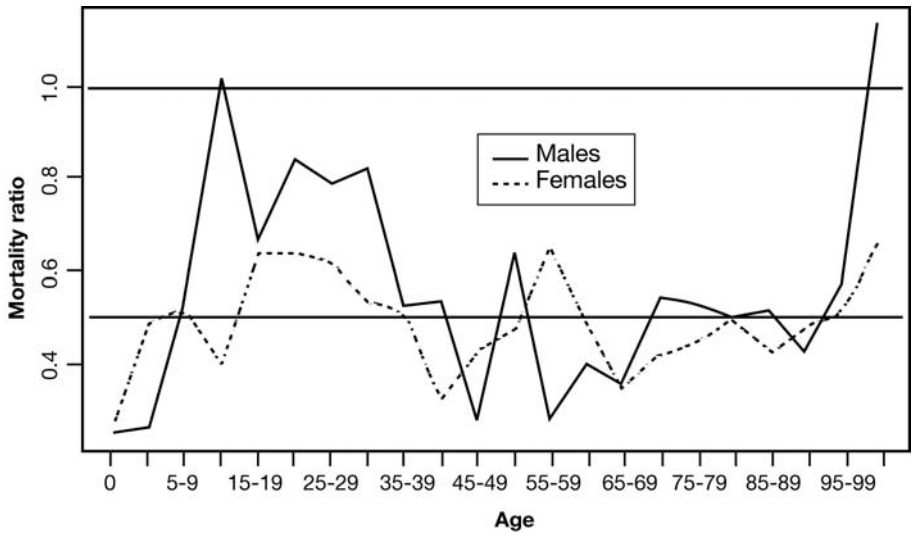

Figure 8.1 Relative mortality of male and female siblings of centenarians compared with birth cohort-matched individuals (controls) from the general American population (survival experience of the controls comes from the Social Security Administration's 1900 birth cohort life table).

The familiality of exceptional longevity

Studying Mormon pedigrees from the Utah Population Database, Kerber and colleagues investigated the impact of family history upon the longevity of 78 994 individuals who achieved at least the age of 65 years (Kerber *et al.*, 2001). The relative risk of survival (λ_s) calculated for siblings of probands achieving the 97th percentile of ''excess longevity'' (for males this corresponded with an age of 95 years, and for women with an age of 97 years) was 2.30. Recurrence risks among more distant relatives in the Mormon pedigrees remained significantly greater than 1.0 for numerous classes of relatives leading to the conclusion that single-gene effects were at play in this survival advantage. The Mormon study findings closely agree with a study of the Icelandic population in which first degree relatives of those living to the 95th percentile of surviving age were also almost twice as likely to live to the 95th percentile of survival compared with controls (Gudmundsson *et al.*, 2000). Both research groups asserted that the range of recurrent relative risks which they observed indicated a substantial genetic component to exceptional longevity.

To further explore the genetic aspects of exceptional old age, we recently studied 444 centenarian pedigrees containing 2092 siblings (Perls *et al.*, 2002*b*). We compared sibling death rates and survival probabilities with national US death rates and survival probabilities according to the Social Security Administration's life table for the cohort born in 1900. Compared with the 1900 birth cohort, the siblings of centenarians maintained a lifelong reduction in risk for death of approximately one half, even up through very old age.

As shown in Figure 8.1, female siblings had death rates at all ages that were about one-half the national level; male siblings had a similar advantage at most ages, though this diminished somewhat during adolescence and young adulthood. The siblings had an average age of death of 76.7 for females and 70.4 for males compared with 58.3 and 51.5 for the general population. Even after accounting for race and education, the net survival advantage of siblings of centenarians was found to be 16 years greater than the general population.

Siblings may share environmental and behavioral factors early in life that have strong effects throughout life. It would make sense that some of these effects are primarily responsible for the shared survival advantage up to middle age. Alternatively, some of these effects might not become evident until after middle age.

Recent evidence of effects of early life conditions on adult morbidity and mortality points to the importance of adopting a life course perspective in studies of chronic morbidity and mortality in later life as well as in investigations of exceptional longevity (Barker, 1998; Blackwell *et al.*, 2001; Costa, 2000; Elford *et al.*, 1991; Elo, 1998; Hall and Peckham, 1997; Kuh and Ben-Shlomo, 1997; Mosley and Gray, 1993). Characteristics of childhood environment are not only associated with morbidity and mortality at middle age, but they have also been found to predict survival to extreme old age (Preston *et al.*, 1998; 2003). Stone analyzed effects of childhood conditions on survival to extreme old age among cohorts born during the late nineteenth century (Stone, 2002). Key factors predicting survival from childhood to age 110+ for these individuals, most of whom were born between 1870 and 1889, were farm residence, presence of both parents in the household, American-born parents, family ownership of its dwelling, residence in a rural area and residence in the non-South; characteristics similar to those that had been previously shown to predict survival to age 85 (Preston *et al.*, 1998; 2003).

In general however, environmental characteristics, such as socioeconomic status, lifestyle, and region of residence, are likely to diverge as siblings grow older. Thus, if the survival advantage of the siblings of centenarians is mainly due to environmental factors, that advantage should decline with age. In contrast, the stability of relative risk for death across a wide age range suggests that the advantage is due more to genetic than to environmental factors.

Whereas death rates reflect the current intensity of death at a moment in time, survival probability reflects the cumulative experience of death up to that moment in a cohort's life history. Thus, a relatively constant advantage from moment to moment (as seen in the relative death rates) translates into an increasing survival advantage over a lifetime (as seen in the relative survival probabilities). This is seen in Table 8.1, which shows the relative survival probabilities of the male and female siblings of centenarians at various ages.

Table 8.1. Relative survival probability (RSP) with 95% confidence intervals (CI) of siblings of centenarians versus US 1900 cohort

	Males			Females		
Age	RSP	Lower 95% CI	Upper 95% CI	RSP	Lower 95% CI	Upper 95% CI
20	1.00	1.00	1.00	1.00	1.00	1.00
25	1.00	0.99	1.01	1.01	1.00	1.02
60	1.18	1.15	1.21	1.12	1.09	1.14
65	1.29	1.25	1.33	1.16	1.13	1.19
70	1.48	1.42	1.53	1.24	1.21	1.28
75	1.68	1.60	1.77	1.36	1.31	1.41
80	2.03	1.90	2.16	1.54	1.47	1.60
85	2.69	2.47	2.91	1.83	1.73	1.93
90	4.08	3.62	4.54	2.56	2.39	2.74
95	8.35	6.98	9.71	4.15	3.73	4.57
100	17.0	10.8	23.1	8.22	6.55	9.90

By the age of 100 years, the relative survival probability for siblings of centenarians is 8.2 for women and 17 for men. From the analysis of death rates, we know that the siblings' survival advantage does not increase as the siblings age. Rather, the siblings' relative probability of survival is a cumulative measure and reflects their lifelong survival advantage over the general population born around the same time. The marked increase in relative survival probability and sustained survival advantage in extreme old age could be consistent with the forces of demographic selection, in which genes or environmental factors (or both) that predispose to longevity win out over those that are associated with premature or average mortality. The substantially higher relative survival probability values for men at older ages might reflect the

fact that male mortality rates are substantially higher than female mortality rates at these ages and, thus, that men gain a greater advantage from beneficial genotypes than women do. Another possibility is that men require an even more rare combination of genetic and environmental factors to achieve extreme age than women do (Perls and Fretts, 1998). Either possibility could explain why men make up only 15% of centenarians.

A proposed multifactorial model for exceptional longevity and exceptional survival phenotypes

The fact that siblings maintain half the mortality risk of their birth cohort from age 20 to extreme age suggests that multiple factors contribute to achieving exceptional longevity. For a simplistic example, sociodemographic advantages may play key roles at younger ages while genetic advantages may distinguish the ability to go from old age to extreme old age. Undoubtedly exceptional longevity is much more complicated, with temporally overlapping roles for major genes, polygenic, environmental and stochastic components. Such a scenario would be consistent with a threshold model, where predisposition for the exceptional longevity trait can be measured on a quantitative scale. Figure 8.2 illustrates the standard threshold model proposed by Falconer (Falconer, 1965) where it is predicted that the proportion of affected relatives will be highest among the most severely affected individuals. In the case of exceptional longevity, perhaps severity can be measured by additional years beyond a certain age (threshold) or in further years of delay in the case of exceptional survival phenotypes, such as age of onset of disability or of certain diseases.

Examples of phenotypes fitting the threshold model are early-onset breast cancer or Alzheimer's disease, where relatives of patients who develop these diseases at unusually young ages, are themselves at increased risk or liability. Thus, a 108 year old's ''liability'' or predisposition for exceptional

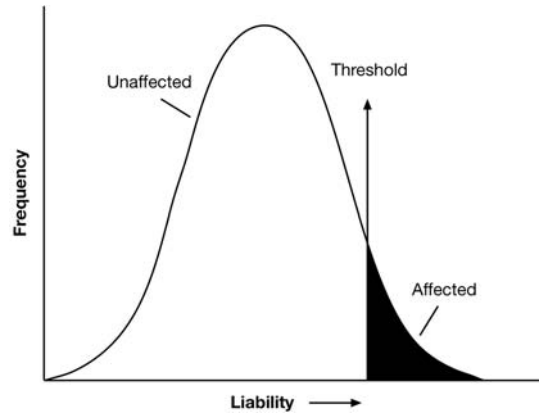

Figure 8.2 Threshold model of a multifactorial trait.

longevity is further beyond the threshold than someone more mildly affected, e.g. a person who died at age 99 years. One interpretation of data indicating the higher relative survival probability of male siblings of centenarians compared with female siblings is that the males carry a higher liability for the trait given the presence of the requisite traits. The model predicts that if a multifactorial trait is more frequent in one sex (as is the case with exceptional longevity which is predominantly represented by females), the liability will be higher for relatives of the less ''susceptible'' sex (males, in the case of exceptional longevity) (Farrer and Cupples, 1998). While we have not yet looked at relative survival probability of siblings of male versus female probands (something that certainly needs to be done), these elevated risks for male versus female siblings are interesting in this context. The model also predicts that the risk for exceptional longevity will be sharply lower for second-degree relatives compared to first-degree relatives, another observation we hope to test by having access to many expanded pedigrees.

Important questions to pursue in light of this proposed model are:

- Are siblings of male centenarians more likely to achieve exceptional longevity themselves compared to siblings of females with exceptional longevity?

- Does this risk for exceptional longevity in siblings and offspring increase with the age of the proband of either sex? What are the quantitative differences of such ''liability'' with or without accounting for the gender of the proband?
- While analysis for reduced risk among second-degree relatives for exceptional longevity is/will be difficult, such analyses may be feasible for exceptional survival phenotypes when comparing cousins in the offspring generation (should we choose to expand data collection to this group).

The ramifications of this model holding true for exceptional longevity (and/or exceptional survival phenotypes) include:

- The older the subject the better the chances of discovering traits predisposing for exceptional longevity.
- There are gender-related differences in both relatives and probands in ''liability'' for exceptional longevity given the presence of specific traits conducive to exceptional longevity.

Genes predisposing to exceptional longevity

It is practically intuitive to state that centenarians outlive those who are relatively predisposed to age-related fatal illnesses and that they are less likely to have environmental and genetic exposures that contribute to death at earlier ages. This selection phenomenon, called demographic selection, is exemplified by the fact that the apolipoprotein E ε4 allele, associated with heart disease and Alzheimer's disease, is rare in centenarians, whereas the prevalence of an alternative allele, ε2, is relatively high (Schachter *et al.*, 1994). Along the same lines, it is likely that there are certain environmental exposures that are rare among centenarians as well, such as tobacco, obesity and bullets. Richard Cutler, in what is now a classic paper in gerontology, proposed that persons who achieve extreme old age do so in great part because they have genetic variations that affect the basic mechanisms of aging and that result in a uniform decreased susceptibility to age-associated diseases (Cutler, 1975). Our studies and those of others researching the oldest old have noted that persons who achieve extreme old age probably lack many of the variations (the ''disease genes'') that substantially increase risk for premature death by predisposing persons to various fatal diseases, both age-associated and non-age-associated (Schachter, 1998). More controversial is the idea that genetic variations might confer protection against the basic mechanisms of aging or age-related illnesses (the ''longevity-enabling genes'') (Perls *et al.*, 2002*a*). The progressive selecting out of more and more genetically fit persons of very old age lays the foundation for a simpler model for sorting out the genetics of aging and longevity.

The discovery of genetic variations that explain even 5% to 10% of the variation in survival to extreme old age could yield important clues about the cellular and biochemical mechanisms that affect basic mechanisms of aging and susceptibility to age-associated diseases. The elevated relative survival probability values found among the siblings of centenarians (Table 8.1) supported the utility of performing genetic studies to determine what genetic region or regions, and ultimately what genetic variations, centenarians and their siblings have in common that confers their survival advantage (McCarthy *et al.*, 1998). Centenarian sibships from the New England Centenarian Study were included in a genome-wide sibling-pair study of 308 persons belonging to 137 families with exceptional longevity. According to nonparametric analysis, significant evidence for linkage was noted for a locus on chromosome 4, near microsatellite D4S1564 (Puca *et al.*, 2001).

The interval on chromosome 4 spanned 12 million base pairs and contained approximately 50 putative genes. In order to identify the specific gene and gene variants impacting lifespan, our genetics colleagues Drs Annibale Puca, Bard Geesaman, Louis Kunkel and Mark Daly performed a haplotype-based fine mapping study of the interval. A detailed haplotype map was created of the chromosome 4 locus that extended over

12 million base pairs and involved the genotyping of over 2000 single nucleotide polymorphism (SNP) markers in 700 centenarians and 700 controls. The study identified a haplotype, approximating the gene microsomal transfer protein (MTP), defined by two SNPs that accounted for all of the statistical distortion detected in the region. Statistically, the result appears to be robust, with a relative risk of nearly 2 ($p < 2 \times 10^{-9}$). With interest narrowing in on a single gene, all known SNP polymorphisms for MTP and its promoter were genotyped in 200 centenarians and 200 controls (young individuals). After haplotype reconstruction of the area was completed, a single haplotype, which was under-represented in the long-lived individuals, accounted for the majority of the statistical distortion at the locus (~15% among the long lived individuals versus 23% in the controls). MTP is a rate limiting step in lipoprotein synthesis and may affect longevity by subtly modulating this pathway. This study supports the feasibility of fine mapping linkage peaks using association studies and the power of using the centenarian genome to identify genes impacting longevity and the diseases of aging (Puca *et al.*, 2001).

Dr Nir Barzilai and his colleagues studying Ashkenazi Jewish centenarians and their families recently found another cardiovascular pathway and gene that is differentiated between centenarians and controls (Barzilai *et al.*, 2003). In Dr Barzilai's study, controls are the spouses of the centenarians' children. It was noted that high-density lipoprotein (HDL) and low-density lipoprotein (LDL) particle sizes were significantly larger among the centenarians and their offspring and the particle size also differentiated between subjects with and without cardiovascular disease, hypertension and metabolic syndrome. In a candidate gene approach the researchers then searched the literature for genes that impact upon HDL and LDL particle size and the came up with hepatic lipase and cholesteryl ester transfer protein (CETP). Comparing centenarians and their offspring against controls, one variation of CETP was noted to be significantly increased among those with or predisposed to exceptional longevity. Given our findings that cardiovascular disease is significantly delayed among the offspring of centenarians and that 88% of centenarians either delay or escape cardiovascular disease and stroke beyond the age of 80 years, it makes sense that the frequency of genetic polymorphisms that play a role in the risk for such diseases (such as Apolipoprotein E ε-4) would be differentiated between long lived individuals and the general population (Rebeck *et al.*, 1994; Schachter *et al.*, 1994; Tilvis *et al.*, 1998; Smith, 2000; van Bockxmeer, 1994).

Conclusion

Studies thus far indicate that exceptional longevity is strongly familial and suggest that a significant component of that familiality is genetic. Once thought to be much too complex a puzzle of environmental, behavioral, genetic and stochastic factors, exceptional longevity and likely associated exceptional survival phenotypes appear to be associated with a number of discernable and influential factors. From age 20 to extreme old age, the siblings of centenarians appear to maintain half the mortality risk of their birth cohort. It is possible that much of the survival advantage at younger ages is related to environmental and behavioral advantages that family members would have in common. Genetic variations that confer a relative survival advantage might confer a selective effect at older ages.

Discovering genes that could impart the ability to live to old age, while compressing the period of disability toward the end of life, would yield important insight into how the aging process increases susceptibility to diseases associated with aging and how this susceptibility might be modulated (Hitt *et al.*, 1999). We anticipate that human longevity-enabling genes will be found to influence aging at its most basic levels, thus affecting a broad spectrum of genetic and cellular pathways in a synchronous manner.

The centenarian genome should also be an efficient tool for ferreting out disease genes. Comparison of single-nucleotide polymorphism frequencies implicated in diseases in centenarians and in persons with these diseases should show clinically relevant polymorphisms. Another approach that researchers are in the early stages of understanding is differential gene expression in models known to slow the aging process, such as caloric restriction (Lee *et al.*, 1999). This may prove to be another potent tool for discovering longevity-enabling genes. The hope, of course, is that these gene discoveries will help identify drug targets and create drugs that would allow persons to become more ''centenarian-like'' by maximizing the period of their lives spent in good health.

The careful phenotyping of numerous animal and human models of aging, the collection of genetic material, and the current explosion in molecular genetics data and techniques are soon likely to fill important gaps in the aging puzzle. Complex gene–gene and gene–environment interactions will certainly complicate our ability to understand how genes affect aging. However, with the power of demographic selection, centenarians have already proven helpful in deciphering some polymorphisms and genetic loci associated or not associated with exceptional old age. The children of centenarians, who seem to be following closely in their parents' footsteps, might yield additional discoveries about phenotypic and genetic correlates of successful aging.

REFERENCES

Barker, D.J. (1998). *Mothers, babies, and health in later life*, London: Churchill Livingstone.

Barzilai, N., Atzmon, G., Schechter, C. *et al.* (2003). Unique lipoprotein phenotype and genotype associated with exceptional longevity. *JAMA*, **290**, 2030–40.

Blackwell, D.L., Hayward, M.D. and Crimmins, E.M. (2001). Does childhood health affect chronic morbidity in later life? *Soc Sci Med*, **52**, 1269–84.

Costa, D.L. (2000). Understanding the twentieth-century decline in chronic conditions among older men. *Demography*, **37**, 53–72.

Crimmins, E. and Ingegneri, D.G. (1993). Trends in health among the American population In Bartlett, D. and Rappaport, A.M. (eds.), *Demography and retirement: The 21st Century*. Westport, CT: Praeger.

Cutler, R.G. (1975). Evolution of human longevity and the genetic complexity governing aging rate. *Proc Natl Acad Sci USA*, **72**, 4664–8.

Elford, J., Whincup, P. and Shaper, A.G. (1991). Early life experience and adult cardiovascular disease: longitudinal and case-control studies. *Int J Epidemiol*, **20**, 833–44.

Elo, I. (1998). Childhood conditions and adult health: evidence from the health and retirement study. In *Population Aging Research Center Working Papers*. Philadelphia, PA.

Evert, J., Lawler, E., Bogan, H. and Perls, T. (2003). Morbidity profiles of centenarians: survivors, delayers, and escapers. *J Gerontol A Biol Sci Med Sci*, **58**, 232–7.

Falconer, D.S. (1965). The inheritance and liability to certain disease estimated from the incidence among relatives. *Annals of Human Genetics*, **29**, 51–76.

Farrer, L. and Cupples, A. (1998). Determining the genetic component of a disease. In Haines, J. and Pericak-Vance, M.A. (eds.), *Approaches to gene mapping in complex human diseases*. NewYork: Wiley-Liss.

Fontaine, K.R., Redden, D.T., Wang, C., Westfall, A.O. and Allison, D.B. (2003). Years of life lost due to obesity. *Jama*, **289**, 187–93.

Fraser, G.E. and Shavlik, D.J. (2001). Ten years of life: Is it a matter of choice? *Arch Intern Med*, **161**, 1645–52.

Freedman, V.S. and Soldo, B.J. (1994). *Trends in disability at older ages: Summary of a workshop*. Washington, DC: National Academy Press.

Fries, J.F. (1980). Aging, natural death, and the compression of morbidity. *N Engl J Med*, **303**, 130–5.

Gruenberg, E. (1977). The failure of success. *Milbank Memorial Fund Quaterly*, **5**, 3–34.

Gudmundsson, H., Gudbjartsson, D.F., Frigge, M., Gulcher, J.R. and Stefansson, K. (2000). Inheritance of human longevity in Iceland. *Eur J Hum Genet*, **8**, 743–9.

Hall, A.J. and Peckham, C.S. (1997). Infections in childhood and pregnancy as a cause of adult disease – methods and examples. *Br Med Bull*, **53**, 10–23.

Herskind, A.M., McGue, M., Holm, N.V. *et al.* (1996). The heritability of human longevity: a population-based

study of 2872 Danish twin pairs born 1870–1900. *Hum Genet*, **97**, 319–23.

Hitt, R., Young-Xu, Y., Silver, M. and Perls, T. (1999). Centenarians: the older you get, the healthier you have been. *Lancet*, **354**, 652.

Kerber, R. A., O'Brien, E., Smith, K. R. and Cawthon, R. M. (2001). Familial excess longevity in Utah genealogies. *J Gerontol A Biol Sci Med Sci*, **56**, B130–9.

Kramer, M. (1983). The increasing prevalence of mental disorders: a pandemic threat. *Psychiatr Q*, **55**, 115–43.

Kuh, D. and Ben-Shlomo, B. (1997). *A life course approach to chronic disease epidemiology*. Oxford: Oxford University Press.

Lee, C. K., Klopp, R. G., Weindruch, R. and Prolla, T. A. (1999). Gene expression profile of aging and its retardation by caloric restriction. *Science*, **285**, 1390–3.

Ljungquist, B., Berg, S., Lanke, J., McClearn, G. E. and Pedersen, N. L. (1998). The effect of genetic factors for longevity: a comparison of identical and fraternal twins in the Swedish Twin Registry. *J Gerontol A Biol Sci Med Sci*, **53**, M441–6.

Manton, K. G., Corder, L. and Stallard, E. (1997). Chronic disability trends in elderly United States populations: 1982–1994. *Proc Natl Acad Sci USA*, **94**, 2593–8.

McCarthy, M. I., Kruglyak, L. and Lander, E. S. (1998). Sib-pair collection strategies for complex diseases. *Genet Epidemiol*, **15**, 317–40.

McGue, M., Vaupel, J. W., Holm, N. and Harvald, B. (1993). Longevity is moderately heritable in a sample of Danish twins born 1870–1880. *J Gerontol*, **48**, B237–44.

Mitchell, B. D., Hsueh, W. C., King, T. M. *et al.* (2001). Heritability of life span in the Old Order Amish. *Am J Med Genet*, **102**, 346–52.

Mosley, W. and Gray, R. (1993). Childhood precursors of adult morbidity and mortality in developing countries: implications for health programs. In Gribble, J. and Preston, S. (eds.), *The Epidemiological transition: policy and planning implications for developing countries*. Washington, DC: National Academy Press, pp. 69–100.

Olshansky, S. J. and Ault, A. B. (1986). The fourth stage of the epidemiologic transition: the age of delayed degenerative diseases. *Milbank Q*, **64**, 355–91.

Perls, T. (1997). Symposium: Cognitive and functional status of centenarians. Reports from four studies. *The Gerontologist*, **37**, 37.

Perls, T. and Fretts, R. (1998). Why women live longer than men. *Sci Amer Pres*, 100–7.

Perls, T., Kunkel, L. M. and Puca, A. A. (2002*a*). The genetics of exceptional human longevity. *J Am Geriatr Soc*, **50**, 359–68.

Perls, T. T., Wilmoth, J., Levenson, R., Drinkwater, M., Cohen, M., Bogan, H., Joyce, E., Brewster, S., Kunkel, L. and Puca, A. (2002*b*). Life-long sustained mortality advantage of siblings of centenarians. *Proc Natl Acad Sci USA*, **99**, 8442–7.

Preston, S., Elo, I. T., Hill, M. E. and Rosenwaike, I. (2003). *The demography of African Americans, 1930–1990*. Dordrecht, Netherlands: Kluwer Academic Publishers.

Preston, S. H., Hill, M. E. and Drevenstedt, G. L. (1998). Childhood conditions that predict survival to advanced ages among African-Americans. *Soc Sci Med*, **47**, 1231–46.

Puca, A. A., Daly, M. J., Brewster, S. J. *et al.* (2001). A genome-wide scan for linkage to human exceptional longevity identifies a locus on chromosome 4. *Proc Natl Acad Sci USA*, **98**, 10505–8.

Rebeck, G. W., Perls, T. T., West, H. L. *et al.* (1994). Reduced apolipoprotein epsilon 4 allele frequency in the oldest old Alzheimer's patients and cognitively normal individuals. *Neurology*, **44**, 1513–16.

Schachter, F. (1998). Causes, effects, and constraints in the genetics of human longevity. *Am J Hum Genet*, **62**, 1008–14.

Schachter, F., Faure-Delanef, L., Guenot, F. *et al.* (1994). Genetic associations with human longevity at the APOE and ACE loci. *Nat Genet*, **6**, 29–32.

Silver, M., Newell, K., Hyman, B. *et al.* (1998). Unraveling the mystery of cognitive changes in old age: correlation of neuropsychological evaluation with neuropathological findings in the extreme old. *Int Psychogeriatr*, **10**, 25–41.

Smith, J. D. (2000). Apolipoprotein E4: an allele associated with many diseases. *Ann Med*, **32**, 118–27.

Stone, L. (2002). In *Annual meeting of the population association of America (Proceedings)*, Atlanta, GA.

Tilvis, R. S., Strandberg, T. E. and Juva, K. (1998). Apolipoprotein E phenotypes, dementia and mortality in a prospective population sample. *J Am Geriatr Soc*, **46**, 712–15.

van Bockxmeer, F. M. (1994). ApoE and ACE genes: impact on human longevity. *Nat Genet*, **6**, 4–15.

Verbrugge, L. M. (1984). Longer life but worsening health? Trends in health and mortality of middle-aged and

older persons. *Milbank Mem Fund Q Health Soc*, **62**, 475–519.

Waidmann, T. and Manton, K. G. (1998). Washington, DC: The Urban Institute.

Wechsler, H., Rigotti, N. A., Gledhill-Hoyt, J. and Lee, H. (1998). Increased levels of cigarette use among college students: a cause for national concern. *Jama*, **280**, 1673–8.

Wei, M., Kampert, J. B., Barlow, C. E. *et al.* (1999). Relationship between low cardiorespiratory fitness and mortality in normal-weight, overweight, and obese men. *Jama*, **282**, 1547–53.

9

The MHC paradigm: genetic variation and complex disease

Adrian P. Kelly and John Trowsdale

The MHC, the region of the genome widely believed to be associated with disease resistance, is in fact linked with more disease susceptibility than any other region of the human genome (Price *et al.*, 1999). One explanation for this paradox is that there is a net cost in providing resistance to infection. In other words, improved resistance to infection, manifest as a more effective immune response, results in a greater propensity to autoimmune disease. In this article we will explore this proposal by examining the main features of the MHC, the functions of the genes it contains and its role in disease (Lechler and Warrens, 2000; Marsh *et al.*, 2000).

Features of the MHC

The human MHC is a gene-dense region that contains genes for the classical class I molecules HLA -A, -B and -C as well as class II molecules DP, DQ and DR, all of which are highly polymorphic (Figure 9.1). The ~4Mbp complex is located on chromosome 6p21.3 and contains over 220 genes. The main known function of MHC class I and class II molecules is to present peptide fragments from pathogens on the cell surface for recognition by T cells. This guides killing of virus-infected cells, activation of phagocytes, or production of specific antibodies. Pathogens are known to employ a variety of different strategies in their attempts to evade presentation (Vossen *et al.*, 2002) but features of MHC class I and class II genes help to counteract evasion. For instance, MHC class I and class II molecules are polygenic and highly polymorphic so that different individuals are likely to select different peptides for presentation. Another striking property of MHC genes, that may also help them to keep pace with rapidly evolving pathogens, is their evolution by sequence exchange, by either intra-allelic or intra-locus gene conversion. The clustering of polymorphic genes in the MHC and strong linkage disequilibrium across the complex may also assist in coordinating the battery of different immune response mechanisms.

The MHC and disease: general considerations

Tiwari and Terasaki's book published in 1985 (Tiwari and Terasaki, 1985), listed hundreds of HLA-associated diseases, even though the first demonstration of a strong association, that of HLA-B27 and ankylosing spondylitis (AS), came only a decade earlier. Meanwhile, many more associations have been added to the list but, apart from Narcolepsy, which shows an almost absolute association with DQB1*0602, none have been as strong as the association of HLA-B27 with AS (Lechler and Warrens, 2000). There are conflicting reports for many of the weaker associations, which include everything from aging and dyslexia to autism and sleepwalking (Lecendreux *et al.*, 2003). In some of these studies insufficient numbers of cases and controls were examined to obtain

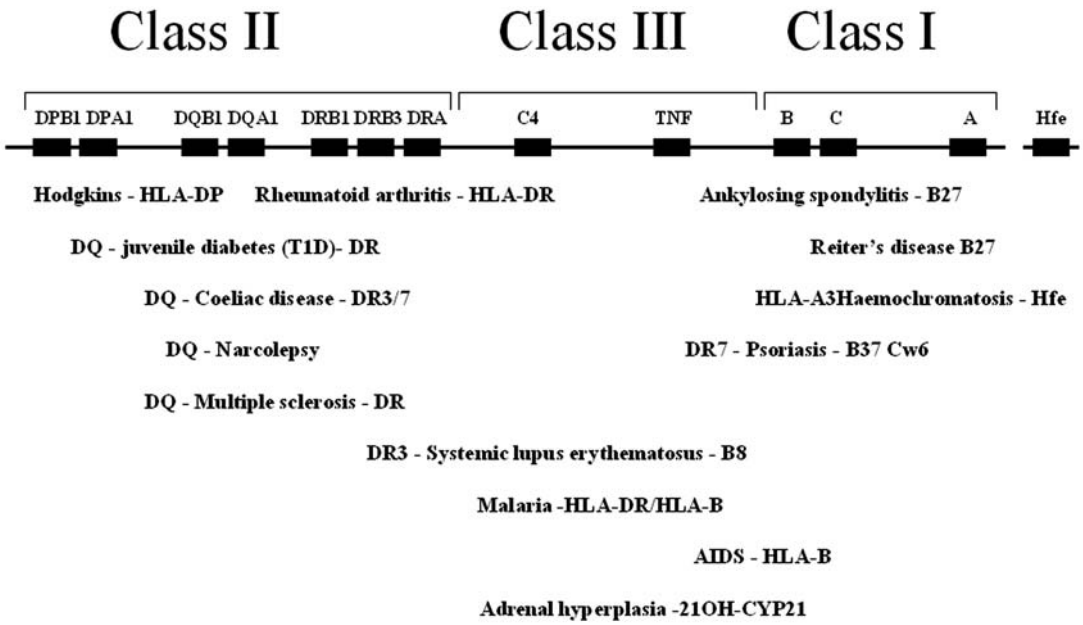

Figure 9.1 MHC-associated diseases are still imprecisely mapped although some of the main HLA class I and class II, as well as additional linked genes, are more strongly implicated in certain conditions.

statistical significance. In others, including multiple sclerosis (MS), association was found in some populations and not others and meta-analysis was necessary to obtain reliable data. It is tempting to reach the conclusion that most human diseases will show at least a weak association with the MHC. Ascertainment bias has been proposed since the HLA region has been studied more thoroughly than any other region of the genome. However, extensive screens of the genome with micro-satellites and single nucleotide polymorphisms have not found a disease region comparable in significance.

A useful step in investigating a genetic association with disease is to study concordance rates in identical, compared to non-identical, twins. This gives an impression of the contribution of the inheritance of genes to the condition. It must be appreciated that in autoimmune diseases environmental effects, such as infections, are profound and figures on the contribution of genetics versus the contribution from the environment must be viewed with this in mind. The sibling risk ratio (λs) is calculated from the risk to sibs divided by the prevalence in the population. For type I diabetes mellitus (T1D) the value is around 15, depending on the population. A λs less than 5 usually means a small genetic contribution.

The strength of association of MHC genotype with disease can be measured by comparing the frequency of different MHC alleles from panels of patients with their frequency in the normal population. The control population must be carefully selected for matching ethnicity, age, gender and environment. It is becoming appreciated that large numbers of cases and controls must be studied and these studies are best performed in collaborations. There are advantages to using non-parametric methods such as allele sharing. The transmission/disequilibrium test (TDT), for example, is carried out with data on transmission of marker alleles from parents heterozygous for the marker to affected offspring (Ewens and Spielman, 1995) (see Chapter 4). In families of affected patients, two siblings with the same

disease are more likely than expected by chance to share the same MHC haplotype and this can be used to measure the contribution of linked genes. Most studies of autoimmune diseases are consistent with significant environmental and genetic components. The genetic component most likely comprises many genes although the most significant of these is the MHC. The others appear to affect immunological tolerance. This has practical significance since once a gene becomes identified as a factor in one autoimmune disease it becomes a strong candidate for other conditions. Very few candidate genes so far appear to encode specific autoantigens. There is evidence of epistasis, at least from mouse models: no single allele is sufficient or necessary to cause disease and the more susceptibility alleles at unlinked loci the greater the risk. Different combinations of susceptibility alleles could contribute to disease in different individuals. For a discussion of analysis of genetic susceptibility, see (Lechler and Warrens, 2000; Vyse and Todd, 1996).

Earlier studies were performed by typing HLA class I alleles serologically, before MHC class II typing was available. Many of these initial associations, with notable exceptions, such as HLA-B27, were subsequently shown to be more strongly associated with MHC class II antigens. The strong linkage disequilibrium (LD) across the MHC region was a key factor in facilitating the initial detection of disease linkage. The extensive LD in the MHC provided an advantage as any polymorphic marker in the region could be used rapidly to obtain evidence of an MHC affect. In the longer term, the LD has been a source of frustration in identifying the key etiologic gene, as was the case with the iron storage disease hemochromatosis (below). For example, the common northern European haplotype A1B8DR3, where HLA-A1, -B8 and -DR3 are in extremely strong LD, is associated with a number of diseases, which were mostly first picked up with HLA-B8 (Price *et al.*, 1999). Only later was HLA-DR3 found to show a better association in some diseases, later refined further to the associated DQ2 locus.

Autoimmune diseases

The MHC is associated with many, if not all, conditions collectively known as autoimmune. These complaints span an extremely broad range of phenotypes but their symptoms all include some form of chronic inflammation. The MHC association does not necessarily invoke an increased susceptibility but may result in faster onset or increased severity. A sample of these conditions is provided in Table 9.1 and is covered more extensively elsewhere (Lechler and Warrens, 2000, Chapter 18 this volume). The diseases are all associated with common HLA allotypes and the vast majority of individuals that express these molecules are free from disease. For example, only ~2% of HLA-B27 individuals exhibit the full symptoms of ankylosing spondylitis (AS). This rises to ~20% in HLA-B27 individuals with affected relatives, consistent with the involvement of additional genes. Most other autoimmune diseases are primarily associated with class II alleles. Even today, the primary associations of most HLA-linked diseases are in doubt and refinement of genetic analysis may lead to variation in a linked DNA sequence. Even so, there is no doubt that MHC class II molecules are major disease determinants but in no case has the precise molecular mechanism been uncovered.

MHC mapping of autoimmune conditions became refined to the extent of identifying precise sequences within a locus that are most closely associated with disease susceptibility. For example, in insulin-dependent diabetes mellitus (now known as type 1 diabetes or T1D), initial linkage to class I led to HLA-DR and subsequently to HLA-DQ. The most abundant amino acid at position 57 of DQB is an aspartic acid, which is able to form a salt bridge across to the α chain at one end of the peptide-binding groove (Janeway *et al.*, 2001). Diabetics in White populations tend to lack the salt bridge and aspartate is replaced by a hydrophobic amino acid. The NOD strain of mouse that develops spontaneous diabetes also lacks the aspartate at position 57 of the orthologous class II molecule A^{g7}.

Table 9.1. Representative MHC-associated diseases

A. *Infectious diseases*	
Disease or agent	*Some associated HLA alleles*[1]
Parasites	
Malaria	B*5301/DRB1*1302 (protection)
Leishmaniasis	Various
Filariasis	Various
Lyme disease	DRB1*0401
viruses	
Epstein–Barr virus	Various
AIDS progression	HLA-B*27/57 (protection) HLA heterozygosity (protection) HLA-B*35
Hepatitis B and C persistence	Various
Hantavirus	HLA-B27
Marek's disease (Chickens)	
Bacteria	
mycobacterial infections	Leprosy (DR2) Tuberculosis (DR2)
B. *Autoimmune diseases and some HLA associations*	
Disease	*Main HLA alleles*[1]
Rheumatoid arthritis	DRB1*0401/0404/0101/0405/1402
Ankylosing spondylitis	HLA-B2702/04/05
Multiple sclerosis	DRB1*1501
Goodpasture's syndrome (renal disease)	DRB1*1501/1502
Graves' disease	DR3
Myasthenia gravis	A1/B8/DR3
Systemic lupus erythematosus	A1/B8/DR3
Hashimoto's thyroiditis	DRB1*0301/DQA1*0501
Pemphigus vulgaris	DRB1*04
Insulin-dependent diabetes mellitus (T1D)	DQB1*0302/0201/04/0303
Celiac disease	DQA1*0501/DQB1*02
Narcolepsy	DQB1*0602
Respiratory diseases	Various
Beryllosis	DPB1-Glu[69]
Eye diseases	Various
Acute anterior uveitis	HLA-B27
Allergy/asthma	Various class II allotypes
C. *Other conditions*	
Haemochromatosis	HLA-HFE
Psoriasis	HLA-C*06
Behcet's disease	HLA-B51
21-0H	CYP21
TAP/Tapasin deficiency IgM	TAPBnull/TAPBPnull
Other class III – associated Hodgkins lymphoma	Weak associations reported
Psychiatric disorders	

*The alleles shown result in increased susceptibility, unless stated otherwise. For references and further details, see (Lechler and Warrens, 2000).

[1] Strongly associated alleles are shown but in many cases several different alleles are associated with resistance or susceptibility to the condition. Different loci can also contribute, for example in T1D where protection is provided by HLA-DQB*0602 and HLA-DRB1*0403 with additional contributions from HLA-DP class I alleles.

Different autoimmune conditions involve different sequences, in different class I or class II molecules, such as HLA-DR in rheumatoid arthritis (RA), -DQ in T1D or HLA-B27 in ankylosing spondylitis (AS). This is not surprising if we consider that the ability of T cells to respond to a particular antigen is specific to MHC genotype. In rheumatoid arthritis (RA), a string of amino acids at residues 70–74 of the DRB chain (the so-called shared epitope: QK/RRAA) is common to most of the susceptibility alleles (Lechler and Warrens, 2000).

In many autoimmune conditions, a particular amino acid sequence most closely associated to the disease has not been precisely defined. Even when a sequence has been identified, as in T1D, this does not account for all of the MHC-associated contribution to disease. Contributions by other MHC genes may be uncovered by studying large numbers of diabetics all of whom carry the same HLA-DQB allele. In view of the clustering of various immunity genes, it is understandable that several different genes in each haplotype could have combinatorial effects. For example, major

contributions from HLA-DQ could be modulated by the other class II loci such as DR and DP. It is not too difficult to see how these molecules, as well as class I allotypes, could modify T cell responses, especially at the level of self tolerance, since the same pool of T cell receptors is involved in T cell selection, either negatively or positively, by all forms of both class I and class II. The roles of other MHC genes may be non-specific. A distinct contribution from different MHC-linked TNF alleles has been proposed in both humans and mice (Jacob *et al.*, 1990). Effects on autoimmunity have also been proposed for the antigen processing (TAP and proteasome) genes embedded in the class II region of the MHC. A discrete cluster of genes in the class III region of the MHC also affect inflammation and their inter-relationship has led to a proposal to name them as MHC class IV genes (Gruen and Weissman, 1997).

Immunopathology

Known mechanisms can involve antibody responses to cell surface or matrix components, for example in Goodpasture's syndrome, where the autoantigen is part of the basement membrane. This disease is organ-specific. In other cases, such as systemic lupus erythematosus (SLE), the disease is systemic and multiple tissues and organs may be affected. Immune complexes are thought to play a major role in the pathology of systemic autoimmune disease. A third group of diseases is characterized by T cell destruction of tissues and associated activated cells. This type includes T1D and RA. The classification of autoimmune diseases as organ specific or systemic reflects the underlying etiology. In some families with organ-specific disease several family members may be affected but with different organs involved. Similarly, systemic autoimmunity can affect multiple individuals in a family or a single individual can have more than one systemic condition, such as SLE together with Sjogren's syndrome.

Environmental effects

The environment plays a major role in determining the onset and course of most autoimmune diseases. These effects contribute up to 40% of the susceptibility (Janeway *et al.*, 2001). Most environmental factors remain to be determined although there has been no shortage of candidates. One reason why they may be so difficult to identify is that the disease symptoms may develop many years after the triggering events in the environment. Top of the list of candidates for the environmental triggers are infections. Intestinal bacteria are obvious candidates for the HLA-B27-associated conditions. Although many different organisms have been proposed as triggers, identification of precise microbial species has proved elusive so far. It may turn out that the triggers include a variety of non-specific infections.

An interesting alternative possibility is that infections act to prevent disease by shaping the T cell repertoire and it is the lack of such infections in modern western cultures that may be the reason for the escalating frequency of autoimmunity. It is proposed that infections during early life could direct the maturing immune system toward TH1, which counterbalance pro-allergic responses of TH2 cells. A reduction in the microbial burden results in weak TH1 dependence and unrestrained TH2 responses that allow an increase in allergy and autoimmunity. This idea is not fully supported by observations that the prevalence of TH1-autoimmune diseases is also increasing (Yazdanbakhsh *et al.*, 2002).

Non-MHC genes are also involved

A case can be made for contributions of several and, most likely, many different genes in autoimmunity, although the strongest contribution appears to be from the MHC. Where other genes have been identified, they are also generally implicated in immune recognition. In SLE for example, genes specifying antigen clearance, tolerance induction or hormonal status are all involved.

Most autoimmune diseases display a marked sex bias.

The levels of variation in these other contributory genes tend to be minor in comparison to HLA. Like MHC genes, the variant alleles are common in the population and do not result in marked deleterious effects expected of a 'mutant' allele. Phenotypic effects invoked by these minor variants are generally subtle and uncovering them demands a panoply of approaches, to obtain and compare data from genetic mapping and phenotyping in both humans and mice. For example, a role for the immunomodulatory co-receptor CTLA-4 has been long suspected from functional considerations as well as genetic data but its mechanism has been elusive. A study recently uncovered effects of DNA sequences distal to the main body of the gene that resulted in altered levels of splice variants that could make soluble molecules, which antagonise binding of the membrane-associated form to its ligand (Ueda *et al.*, 2003).

The effects of other loci are not necessarily simply additive. Recent studies have searched for the involvement of KIR molecules in MHC-associated disease since, like HLA, these genes are highly polymorphic and they interact with MHC class I molecules as ligands (Trowsdale, 2001). KIR genes are encoded on chromosome 19 and are inherited independently of the MHC. This coupling of two sets of polymorphic receptors in immune recognition has suggested studies of combinations of interacting HLA and KIR alleles. Subtle epistatic effects have been noted in common disorders (Martin *et al.*, 2002*a*; 2002*b*). Not unexpectedly, some of the most marked effects have been in transplantation, where combinations of NK cells which have already been highly selected in a host are juxtaposed with target tissues from another individual (Parham and McQueen, 2003). The most dramatic effects so far suggest that these normally deleterious responses could be harnessed by selected mis-matching to provide a beneficial graft versus leukemia response (Karre, 2002).

Mechanisms for MHC associations

Several plausible mechanisms have been proposed to explain the MHC association with autoimmunity but in no case has the precise mechanism been identified. Identification of the mechanism involved in autoimmunity is confounded by the problem that the symptoms of disease tend to occur long after the initiating events. By the time the condition has been diagnosed the inflammation around the target tissue has stimulated recruitment of T cells to a wide range of antigens. One popular idea is the molecular mimicry hypothesis. It is proposed that T cells activated by presentation of a microbial antigen cross-react with peptides present in a self antigen. The associations of particular MHC alleles arise because these would be the allotypes that present the microbial peptide or the cross-reactive self peptide.

An alternative view is that the association with a particular MHC molecule is indirect and reflects tolerisation of the T cell repertoire to self antigens. In other words, self peptides may drive the positive or negative selection of different clones of developing thymocytes. In many diseases a particular target cell type has been identified and in some the actual protein, or autoantigen, within that cell is known. In T1D, for example, the autoimmune T cell response is directed at pancreatic β cells that make insulin hormone. Insulin is an autoantigen but several other proteins in these cells are also targeted.

Although a mechanism invoking an effect of HLA molecules on tolerance is compelling there are other possibilities. One issue that is difficult to rule out completely is of a cryptic infection (Pleister and Eckels, 2003). The chronic presence of a virus that so far has evaded detection could be expected to promote inflammation and autoimmune sequelae. One of the classic MHC associations, that of HLA-B27 and AS, may also be independent of classic tolerance repertoire mechanisms. Disease symptoms can be induced in rats by transgenesis of HLA-B27 in the absence of β2m. There is evidence that the B27 molecule is prone to form an unusual

structure that may mimic that of class II and could stimulate CD4 T cells aberrantly (Allen *et al.*, 1999). Another interpretation of the data is that HLA-B27 results in β2m deposition (Uchanska-Ziegler and Ziegler, 2003).

In summary, the current view of autoimmune disease involves a predisposition resulting from a combination of genetic and environmental effects. HLA provides the major genetic contribution in addition to variation in many other immunity genes. Amongst the individuals with a genetic predisposition, only a small minority exhibit frank disease. This may be triggered in those people by immune response to infection during which some T cell clones escape from tolerance and become reactive against self peptides. The inflammation and destruction over time broadens the immune response and leads to further uncontrolled inflammation and impairment of physiological functions of the tissue.

Infectious diseases

There is strong evidence for association of only a few infectious diseases with the MHC (Table 9.1, Figure 9.1) although infection is believed to drive MHC variation (see Chapter 18). It has been proposed that the emphasis on autoimmunity is largely because research funding for studying these diseases, prevalent in western populations, has greatly exceeded that for infections that are more usually consigned to poorer countries. In fact, the funding emphasis has shifted considerably since the emergence of modern infectious diseases such as AIDS and SARS.

The large number of HLA class I and class II alleles is thought to be maintained in the population by some form of balancing selection. Some infectious organisms have evolved elaborate mechanisms to avoid immune recognition. Many of these mechanisms involve HLA antigens, directly or indirectly, by either downregulating MHC expression in the infected cell or by remodeling protein antigens so that peptides no longer bind to the host HLA molecules. Statistically, these escape variants tend to appear most often for those HLA allotypes that are most frequent in the infected population. Over time, in response to the infection, these allotypes may become less common in that population and eventually new variants of the pathogen may emerge. This pattern of adaptations both by the pathogen and by distribution of allotypes in host populations leads to a dynamic, unpredictable relationship. HLA antigens associated with infections therefore can differ markedly in different populations. In contrast, autoimmune diseases tend to be associated with similar MHC allotypes in different populations, although this is not always the case. The epidemiological model that best fits MHC selection for disease resistance is frequency-dependent selection, where rare types are at an advantage. A different, though not mutually exclusive, model for describing the relationship of HLA alleles with infections is that of heterozygous advantage. Put simply, this means that pathogens will have to work twice as hard to escape immune recognition in individuals with two different alleles at any one MHC locus, say HLA-B, than they would in a homozygous individual. Either model illustrates how adaptation to pathogens is a force for driving the selection of new HLA alleles.

In AIDS, various studies point to association of different HLA class I alleles with protection or susceptibility. It is now well established that HLA-B27 and HLA-B57 are both associated with slower progression to AIDS after HIV infection. These allotypes may bind peptides that the virus has to maintain intact in order to remain effective. Conversely, various subtypes of HLA-B35 appear to be associated with more rapid progression. A case has been made also for heterozygous advantage, measured as time to onset of full-blown disease after seropositivity. It is logical to propose that HIV escape variants will take longer to develop in an individual heterozygous at HLA -A, -B and -C than in someone who is homozygous. A large study of AIDS patients in western Australia obtained evidence for relentless loss of peptide epitopes binding to host class I allotypes encoded by virus in

that specific host (Moore *et al.*, 2002). This is not surprising since specific T cell clones in AIDS patients have been shown, by use of MHC class I tetramer binding, to be greatly expanded. Recent data are consistent with HIV using host systems to maximise variation as the virus develops (Neuberger *et al.*, 2003).

Resistance to malaria was associated with HLA-B53 in west Africa but a similar association was not evident in east Africa (Hill *et al.*, 1992). A plausible explanation is that strains of the parasite in the two different regions have adapted and those in the eastern part of the country do not provide protective epitopes which can be presented by HLA-B53. Similarly, a complex association of alleles of MHC class I genes, such as HLA-A11 with Epstein Barr virus (EBV) is consistent with loss of peptide epitopes for that particular allotype (de Campos-Lima *et al.*, 1993).

Studies of South American Indians revealed novel HLA-B alleles that supposedly arose in response to regional parasite load. Related populations in North America had maintained unaltered ancestral alleles of HLA-B (Belich *et al.*, 1992). Inspection of the new HLA-B sequences revealed that they could have arisen by intra- or inter-allele gene conversion. This work indicates the way in which MHC class I variation may respond to infectious agents, although these agents remain unidentified.

In summary, the current view is that MHC molecules play a critical role in orchestrating appropriate defence mechanisms in infectious disease. Infectious organisms, such as herpes viruses, go to extraordinary lengths to escape presentation by MHC molecules and resistance to infection is thought to drive the exquisite MHC variation. Resistance or susceptibility to infection is not always associated with a particular MHC allotype since the advantage to the host population is in variety and novelty.

Other MHC-associated diseases

Rare defects in TAP transporters encoded in the MHC result in absence of expression of MHC class I molecules (Gadola *et al.*, 2000). These patients exhibit an interesting immunodeficiency which results in bacterial infections, particularly Pseudomonas infections of the lung. They can survive for many years without identified symptoms but, usually in the third decade of life, may be characterised by a form of granulomatosis.

A number of ''third-party'' diseases are associated with the MHC but are apparently structurally and functionally unrelated to HLA class I and class II loci and associated immune response pathways. They are most likely the result of chance events in evolution that brought genes into juxtaposition with HLA loci, although a functional connection cannot be ruled out. A classic example is the linkage of congenital adrenal hyperplasia, which is due to defective alleles in the CYP21 genes that encode an enzyme involved in biosynthesis of steroids, by cortical cells in the adrenal gland, namely 21-hydroxylase.

The initial association of HLA-A3 with hemochromatosis, a common iron storage condition, was unmistakable and this led most researchers to search in vain for a molecular mechanism involving HLA-A3, or at least polymorphism in a tightly associated gene. Eventually, when painstaking haplotype analysis was performed, which involved the construction of multiple single chromosome 6 somatic cell hybrids, the key haemochromatosis locus was identified, ~4Mbp telomeric of HLA-A3 (Feder *et al.*, 1996). It turned out that the HFE gene encodes a class I-like heavy chain, which couples with β2-microglobulin and regulates iron uptake by association with transferrin receptor. The allele of HFE defective in this function is in strong linkage disequilibrium with HLA-A3.

Psoriasis has been associated with HLA-C*06 and was assumed to be an autoimmune condition. The finding of a set of genes expressed in skin and tightly linked to HLA-C has caused this hypothesis to be revisited. Extensive genetic sleuthing and even complete sequencing of large sections of haplotypes has failed to identify the key gene affecting psoriasis in the complex near HLA-C. The uncomfortable if most parsimonious

explanation to date is that the disease is affected by a complex set of genes on a haplotype (Veal *et al.*, 2002).

Genetic susceptibility to Abacavir hypersensitivity, a potent drug used in anti-HIV therapy, is carried on the ancestral haplotype HLA-B*5701, HLA-DR7, and HLA-DQ3. As mentioned above, HLA-B57 is associated with long-term non-progression of the disease (Migueles *et al.*, 2000). This finding has necessitated careful screening of patients for sensitivity before use of the drug. The gene specifying HIV and resistance to the drug may map to the MHC class III region (Mallal *et al.*, 2002).

In summary, each additional disease associated with the MHC invokes a unique set of circumstances. This is not surprising in view of the 220 genes tightly packed into this complex, which represents only 0.1% of the genome but up to 1% of the expressed genes. There may be some common factors which have helped to maintain linkage of these genes in haplotypes, such as the set of genes implicated in psoriasis.

The price of disease resistance

As pointed out above, resistance to infection appears to be the *raison d'etre* of multiple linked MHC genes, which may act in concert. At first sight it is odd that this region is associated with more disease susceptibility than any other region of the human genome. Is there any direct evidence that improved resistance to infection, and a more effective immune response, means a greater chance of autoimmune disease? There are precedents from other disease systems, such as the relationship between sickle cell anaemia and resistance to malaria (Allison, 1954). Potentially there is a balance between resistance to bacterial infection and susceptibility to autoimmunity due to FcR variation in mice (Pritchard *et al.*, 2000). There is also direct evidence from in vivo studies that pathogens protect from autoimmunity (Zaccone *et al.*, 2003).

It appears that rapid evolution of multiple defence genes, by duplication and variation, capable of dealing with all potential pathogens comes with a penalty. There is an increased risk that, in a highly complex system such as the immune network, certain combinations of products of different alleles may shift the balance in favour of destructive recognition of self tissues.

REFERENCES

Allen, R. L., O'Callaghan, C. A., McMichael, A. J. and Bowness, P. (1999). Cutting edge: HLA-B27 can form a novel beta 2-microglobulin-free heavy chain homodimer structure. *J Immunol*, **162**, 5045–8.

Allison, A. C. (1954). The distribution of the sickle-cell trait in East Africa and elsewhere, and its apparent relationship to the incidence of subtertian malaria. *Trans Roy Soc Trop Med Hyg*, **48**, 312–18.

Belich, M. P., Madrigal, J. A., Hildebrand, W. H. *et al.* (1992). Unusual HLA-B alleles in two tribes of Brazilian Indians. *Nature*, **357**, 326–9.

de Campos-Lima, P. O., Gavioli, R., Zhang, Q. J. *et al.* (1993). HLA-A11 epitope loss isolates of Epstein–Barr virus from a highly A11+ population. *Science*, **260**, 98–100.

Ewens, W. J. and Spielman, S. S. (1995).The transmission/disequilibrium test: history, subdivision, and admixture. *Am J Hum Genet*, **57**, 455–64.

Feder, J. N., Gnirke, A., Thomas, W. *et al.* (1996). A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. *Nature Genetics*, **13**, 399–403.

Gadola, S. D., Moins-Teisserenc, H., Trowsdale, J., Gross, W. and Cerundolo, V. (2000). TAP deficiency syndrome. *Clin and Exp Immunol*, **121**, 173–8.

Gruen, J. R. and Weissman, S. M. (1997). Evolving views of the MHC. *Blood*, **90**, 4252–65.

Hill, A. V. S., Elvin, J., Willis, A. C. *et al.* (1992). Molecular analysis of the association of HLA-B53 and resistance to severe malaria. *Nature*, **360**, 434–9.

Jacob, C. O., Fronek, Z., Lewis, G. D. *et al.* (1990). Heritable major histocompatibility complex class II-associated differences in production of tumor necrosis factor alpha: relevance to genetic predisposition to systemic

lupus erythematosus. *Proc Natl Acad Sci USA*, **87**, 1233–7.

Janeway, C. A., Travers, P., Walport, M. and Shlomchik, M. (2001). *Immunobiology*. New York: Garland.

Karre, K. (2002). A perfect mismatch. *Science*, **295**, 2029–31.

Lecendreux, M., Bassetti, C., Dauviliiers, Y. *et al.* (2003). HLA and genetic susceptibility to sleepwalking. *Mol Psychiatry*, **8**, 114–17.

Lechler, R. and Warrens, A. (2000). HLA in health and disease. London: Academic Press.

Mallal, S., Nolan, D., Witt, C. *et al.* (2002). Association between presence of HLA-B*5701, HLA-DR7, and HLA-DQ3 and hypersensitivity to HIV-1 reverse-transcriptase inhibitor abacavir. *Lancet*, **359**, 727–32.

Marsh, S. G. E., Parham, P. and Barber, L. D. (2000). *The HLA factsbook*. London: Academic Press.

Martin, M. P., Gao, X., Lee, J. H. *et al.* (2002*a*). Epistatic interaction between KIR3DS1 and HLA-B delays the progression to AIDS. *Nat Genet*, **31**, 429–34.

Martin, M. P., Nelson, G., Lee, J. H. *et al.* (2002*b*). Cutting edge: susceptibility to psoriatic arthritis: influence of activating killer Ig-like receptor genes in the absence of specific HLA-C alleles. *J Immunol*, **169**, 2818–22.

Migueles, S. A., Sabbaghian, M. S., Shupert, W. L. *et al.* (2000). HLA B*5701 is highly associated with restriction of virus replication in a subgroup of HIV-infected long term nonprogressors. *Proc Natl Acad Sci USA*, **97**, 2709–14.

Moore, C. B., John, M., James, I. R. *et al.* (2002). Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. *Science*, **296**, 1439–43.

Neuberger, M. S., Harris, R. S., Di Noia, J. and Petersen-Mahrt, S. K. (2003). Immunity through DNA deamination. *Trends Biochem Sci*, **28**, 305–12.

Parham, P. and McQueen, K. L. (2003). Alloreactive killer cells: hindrance and help for haematopoietic transplants. *Nat Rev Immunol*, **3**, 108–22.

Pleister, A. and Eckels, D. D. (2003). Cryptic infection and autoimmunity. *Autoimmun Rev*, **2**, 126–32.

Price, P., Witt, C., Allcock, R. *et al.* (1999). The genetic basis for the association of the 8.1 ancestral haplotype (A1, B8, DR3) with multiple immunopathological diseases. *Immunol Rev*, **167**, 257–74.

Pritchard, N. R., Cutler, A. J., Uribe, S. *et al.* (2000). Autoimmune-prone mice share a promoter haplotype associated with reduced expression and function of the Fc receptor FcgammaRII. *Curr Biol*, **10**, 227–30.

Tiwari, J. L. and Terasaki, P. I. (1985). *HLA and disease*. New York: Springer-Verlag.

Trowsdale, J. (2001). Genetic and functional relationships between MHC and NK receptor genes. *Immunity*, **15**, 363–74.

Uchanska-Ziegler, B. and Ziegler, A. (2003). Ankylosing spondylitis: a beta2m-deposition disease? *Trends Immunol*, **24**, 73–6.

Ueda, H., Howson, J. M., Esposito, L. *et al.* (2003). Association of the T-cell regulatory gene CTLA4 with susceptibility to autoimmune disease. *Nature*, **423**, 506–11.

Veal, C. D., Capon, F., Allen, M. H. *et al.* (2002). Family-based analysis using a dense single-nucleotide polymorphism-based map defines genetic variation at PSORS1, the major psoriasis-susceptibility locus. *Am J Hum Genet*, **71**, 554–64.

Vossen, M. T., Westerhout, E. M., Soderberg-Naucler, C. and Wiertz, E. J. (2002). Viral immune evasion: a masterpiece of evolution. *Immunogenetics*, **54**, 527–42.

Vyse, T. J. and Todd, J. A. (1996). Genetic analysis of autoimmune disease. *Cell*, **85**, 311–18.

Yazdanbakhsh, M., Kremsner, P. G. and van Ree, R. (2002). Allergy, parasites, and the hygiene hypothesis. *Science*, **296**, 490–4.

Zaccone, P., Fehervari, Z., Jones, F. M. *et al.* (2003). Schistosoma mansoni antigens modulate the activity of the innate immune response and prevent onset of type 1 diabetes. *Eur J Immunol*, **33**, 1439–49.

10

Lessons from single gene disorders

Nicholas D. Hastie

Introduction

It has been estimated that there are on the order of 6000 single gene disorders, most inherited as X-linked recessive, autosomal dominant or autosomal recessive Mendelian conditions (McKusick, 1998). In total, approximately 1% of the population is born with or will develop a disease through carrying such single gene mutations.

Even before the completion of the Human Genome sequence (International Human Genome Sequencing Consortium, 2001; Venter *et al.*, 2001), heroic positional cloning efforts had successfully identified the genes mutated in 1000 or so of these disorders. Since the genome sequence has been available, the remainder are being picked off at an astonishing rate, the limiting factor now being the small number of affected individuals in some of the most rare conditions.

Identification of the genes mutated in Mendelian disorders has led to profound insights into disease mechanisms and, in some instances, has already had a clinical impact. Perhaps the most dramatic examples are phenylketonuria and bowel cancer, where identification of the underlying genetic cause has led to cures through nutritional and surgical intervention, respectively. In a number of cases, identification of the genes has enabled prenatal diagnosis with the offer of termination of pregnancy. Finally, in a few conditions the genetic breakthroughs have led to novel therapies, either pharmacological or immunological. Another enormous bonus is the profound impact that knowledge of these genes and their products has had on our understanding of basic biological processes. For example, groundbreaking biochemical studies on the product of the gene mutated in retinoblastoma have given key insights into the regulation of the cell cycle (Bartek *et al.*, 1996). The APC protein, encoded by the gene mutated in familial adenomatous polyposis (FAP), has been shown to play a pivotal role in the signalling pathways regulated by the Wnt proteins (Fodde *et al.*, 2001). Studies on dystrophin, the protein responsible for Duchenne and Becker muscular dystrophy, have revealed a large protein complex that is essential for the integrity of the muscle membrane and which links the extracellular matrix to the actin cytoskeleton of the muscle cell (Dalkilic and Kunkel, 2003). There are hundreds of such inspiring examples.

In this short review, I will not try to make a comprehensive list of monogenic disorders and their biology. Instead, I will consider the ways in which knowledge of single gene disorders will have a bearing on our attempts to identify and comprehend genetic risk factors in common, multifactorial disease. One of the main conclusions of course is that most single gene disorders are not simple at all – exhibiting locus and allelic heterogeneity, incomplete penetrance, or expressivity. Moreover, the phenotypes are often modified by other genes, the environment or chance events. I will discuss these topics and the alarm bells they may ring when considering multifactorial disease. On the positive side, I will consider how the study of single

gene subsets of complex disorders is providing key insights into common disease mechanisms and in some cases has led to new drugs for common forms of diseases.

Locus and allelic heterogeneity: the good and the bad

The strategies for hunting genetic risk factors in common disease are mainly predicated on the common disease/common variant hypothesis. Although we and others have challenged some of the assumptions inherent in this hypothesis (e.g. Wright and Hastie, 2001), there is no doubt that it applies in a number of cases. The most notable example is the increased risk of early-onset Alzheimer's disease for carriers of the apolipoprotein E ε4 allele and, conversely, the protective effect of the ε2 allele (Roses, 1996).

One daunting theme emerging from single gene disorders which may have considerable and alarming implications for common disease is the (perhaps) surprising level of locus and allelic heterogeneity. Furthermore, such heterogeneity has been recorded in numerous late-onset Mendelian disorders, some of which are subsets of more common, polygenic diseases. In regard to locus heterogeneity, mutations in at least three different genes may lead to early-onset familial Alzheimer's disease (Tanzi and Bertram, 2005; Chapter 29, this volume). Far more impressive perhaps are the 36 different loci so far associated with Mendelian forms of non-syndromic retinitis pigmentosa (RP) and the 48 loci associated with syndromic RP (Wright and Hastie, 2001; Chapter 32, this volume).

With regard to allelic heterogeneity, more than 480 different mutations have been identified in the *BRCA1* gene leading to increased risk of breast and ovarian cancer (Wright and Hastie, 2001). Moreover, more than 730 different mutations have been documented in the LDL receptor gene leading to increased risk of coronary artery disease (Wright and Hastie, 2001).

If such locus and allelic heterogeneity applies to common diseases, in general, this will severely hinder the ability to identify genetic risk factors using (single nucleotide polymorphism) SNP-based association studies. Moreover, if such locus and allelic heterogeneity does apply to common diseases, we will end up classifying them into much more rare genotypic and phenotypic subsets, interacting in different ways with different environmental triggers. This will no doubt have a considerable impact on the pharmaceutical companies' attitude to, and strategies for developing, drugs based on this knowledge. Another implication is that predicting individual risk on the basis of genetic typing will become even more difficult from a statistical point of view.

Whilst acknowledging the downside to all this locus heterogeneity, we cannot forget the positive side. In many cases the different genes, mutated in a particular condition, have been shown to work in common biological or biochemical pathways. On the one hand prior knowledge of a pathway can lead to an intelligent prediction of additional candidate genes once a first gene in a pathway has been associated with a particular condition. On the other hand, identification of different genes mutated in different families with the same disease may reveal new pathways. Clearly identification of key pathways provides great mechanistic insights but also pinpoints the pathway that is crucial in relation to drug development. One of the most striking of many such examples is Alzheimer's disease, which will be discussed in more detail below.

Penetrance, expressivity and genetic modifiers

One of the major benefits expected to derive from genetic studies of common disease, will be the ability to identify individuals at highest risk of developing a particular disease through genetic profiling. For example, from a rigorous theoretical analysis, it has been calculated that it should be

possible to identify the 10% of women who have a 20–30% risk of breast cancer, compared with a 5% risk for women across the whole population (Pharoah *et al.*, 2002; Chapter 15, this volume). This will depend on the success of identifying the genetic factors contributing to at least half of the overall genetic susceptibility to the disease. This seems reasonable, although of course it may still be bedevilled by the existence of many genes with tiny effects or heterogeneity that might involve rare alleles. Moreover, the situation will be complicated by the as yet unknown nature of interactions between these low penetrance genes – whether additive, multiplicative, or epistatic. We do not know how this will work out and very large sample sizes will be required to clarify the situation.

Some companies, misguidedly, are already offering the general public genetic profiling to determine whether they are at increased risk of, say, heart disease. Clearly even if the hunt for genetic risk factors is successful and one can account for most of the genetic contribution to susceptibility, the risk estimates will be broad and apply to the population but what will it mean for an individual? Again we only have to look at single gene disorders to get an inkling of the much greater problems we will find in studying common diseases. There are very few single gene disorders that are completely penetrant, since the outcome is commonly influenced by genetic background and environmental factors. Even single gene disorders which are completely penetrant may show variation in expressivity, again dependent on unknown genetic and environmental factors.

Different mutations within the same gene can of course have very different phenotypic consequences. If the common disease/common variant hypothesis holds up, this may not be much of an issue in common disease but it will be surprising if there is not some allelic heterogeneity. Alternatively, the same mutation can lead to very different outcomes in different members of the same family or different families. The situation may be even more complicated in a recessive condition where different compound heterozygous combinations may have different outcomes. Here I will consider a few specific examples; two dominant Mendelian conditions, hereditary non-polyposis colorectal carcinoma (HNPCC) and Huntington's Disease (HD), and a common recessive condition, cystic fibrosis (CF).

HNPCC arises through mutation in one of several genes encoding components of the DNA mismatch repair pathway, particularly the *MLH1* and *MSH2* genes. Whereas 80% of males carrying the mutation develop bowel cancer, the penetrance is only 30% in females where oestrogens may play a protective role (Farrington and Dunlop, 2004; Chapter 17, this volume). On the other hand, females are at increased risk of endometrial and cervical cancer. Clearly one should be on the look out for such cases of sexual dimorphism in more common cancers with complex etiology or in other common diseases.

HD is an autosomal dominant condition arising through an expanded stretch of polyglutamine, encoding CAG repeats in the Huntington gene (Bates *et al.*, 2002). Individuals with alleles containing fewer than 34 CAG repeats never develop symptoms. However, individuals with >40 repeats will always develop HD. The age of onset for HD varies over a 40-year time period and, remarkably, this correlates inversely with the length of the repeat. Hence it is not surprising that HD is often regarded as an extreme case of genetic determinism and one that has perhaps distorted people's views of the potential predictive power of genetics. However, the expressivity of even HD is influenced by additional genetic and environmental factors. A comprehensive and thorough study of Venezuelan kindreds (The US–Venezuela Collaborative Research Project, 2004) showed that the length of the repeat only accounts for 40–70% of this variable age of onset. Forty percent of the remaining variation is explained by unknown genetic factors, the rest by unknown environmental factors.

CF is the most common lethal autosomal recessive childhood disorder in White populations, occurring in approximately 1:2500 live births

(Zielenski and Tsui, 1995). Patients with CF have abnormal chloride conductance across the apical membrane of epithelial cells, leading to chronic pulmonary obstruction, bacterial colonization of the airways, pancreatic enzyme insufficiency and reduced fertility in males. The defect is due to mutations in the *CFTR* gene which encodes a protein of 1480 amino acids and functions as a chloride channel (Zielenski and Tsui, 1995). Although one mutation, ΔF508, predominates and is found in 68% of CF chromosomes worldwide, over 500 different mutations have been documented. The disease spectrum of CF is highly heterogeneous among the different patients. Extensive reviews of case reports show that it is possible to make broad genotype–phenotype correlations. However, the only consistent phenotypic association with CF genotype is pancreatic sufficiency versus insufficiency (The Cystic Fibrosis Genotype–Phenotype Consortium, 1993). The molecular lesions can be grouped into 5 categories, classes 1–3 classified as severe, and classes 4–5 as less severe. In general, patients with pancreatic insufficiency are homozygotes or compound heterzygotes for two severe mutations, whereas patients with pancreatic sufficiency have at least one mild allele. Other CF symptoms are either indistinguishable among different genotype groups or highly variable, even among sibs with identical mutations.

Infertility in male CF patients is due to obstructive azoospermia secondary to bilateral absence of the vas deferens. A similar autosomal recessive condition with no other CF manifestions is classified as congenital bilateral absence of vas deferens (CBAVD). Remarkably, 50–65% of patients with CBAVD have at least one mutated *CFTR* allele, whereas 10% have two mutated alleles. Furthermore, patients with six other milder pulmonary diseases that are subsets of the CF spectrum have been shown to have at least one *CFTR* mutation (Zielenski and Tsui, 1995). We must conclude that these phenotypes result from a dominant effect of the *CFTR* mutation, or that there is an undected mutation in the second allele – perhaps regulatory – or a mutation in another unlinked gene or genes which modifies the effect of a single allelic mutation. Taken together, all this demonstrates the remarkable heterogeneity of phenotypes arising through different combinations of *CFTR* mutations.

The difficult hunt for modifier loci

Although different *CFTR* mutations are associated with different phenotypic outcomes, *CFTR* allelic heterogeneity itself only partly explains the wide variety of clinical presentation. There is clinical variation in patients with the same *CFTR* mutations and even in members of the same family carrying the same *CFTR* mutations. The assumption is that there are unlinked genetic modifiers of the *CFTR* mutations. Clearly if these genetic modifiers could be identified, they could shed new light on the biology of the system and help to provide more accurate diagnosis and treatment. However, it has proved to be very difficult to identify these additional genetic factors due to genetic heterogeneity and inadequate sample size, both features which are likely to have hampered common disease genetics. Help in identifying *CFTR* genetic modifiers appears to have come from studying mouse models. *CFTR*-deficient mice created by gene targeting developed severe intestinal obstruction leading to death within a few weeks of life. However, the severity of the phenotype and survival varied depending on the genetic background of the mice carrying the *CFTR* mutation. This difference in survival between strains was exploited to map a genetic modifier to mouse chromosome 7 (Rozmahel *et al.*, 1996). Electrophysiological studies showed that the mice with prolonged survival had upregulation of a calcium activated chloride channel and that this might account for the rescue. Drawing on these findings in mice, the same group was able to map a genetic modifier locus in humans to the region of chromosome 19 with conserved synteny to the region of mouse chromosome 7, harboring the modified locus (Zielenski *et al.*, 1999). Although the mouse and human papers were published in

1996 and 1999 respectively, the causative genetic change associated with the modifying activity has still not been pinpointed. Again this experience with something as relatively simple and penetrant as CF shows the magnitude of the task to identify interacting genetic loci in common multifactorial disease.

Mutational spectra: how many types of functional variant?

There is much debate about the likely nature and spectrum of functional variants conferring risk or protection against common disease. Is there any reason to believe that the spectrum of mutations will differ from that observed in single gene disorders? In Mendelian diseases, in-frame base pair substitutions including those that produce nonsense codons are the most frequent, accounting for 59% of all disease-associated mutations. Chromosome deletions account for 22% and insertions/duplications for 7% of all the mutations. In total, regulatory mutations affecting the level of expression account for only 1% of mutations (Krawczak *et al.*, 2000). In most single gene disorders the mutations are likely to have a severe impact on the expression level, structure or function of the gene product. In the vast majority of cases the mutations lead to loss or reduction in activity; there are few examples of true neomorphic or gain-of-function mutations.

Clearly, as discussed below, for any particular single gene disorder there may be a wide range of phenotypic severity and this may be related to different types of mutation in the same gene. Hence for β-thalassemias, the most studied of all human genetic diseases, milder alleles are usually due to missense mutations or single base changes in the promoter region that lead to a reduction in expression levels. A distinct possibility is that functional variants in common disease are more likely to have subtle effects on the levels of gene expression or function equivalent perhaps to the milder thalassemia alleles. At present we just do not have enough examples of rigorously verified functional variants in common disease or quantitative trait loci to make any predictions.

Thalassemias have provided a remarkable paradigm for documenting the mechanistic diversity of mutations which may lead to (more or less) the same disease (Weatherall, 2001; Chapter 21, this volume). Over 200 different β-globin gene mutations have been described in β-thalassemia. In severe β-thalassemia, the most common mutations are splice site, frame shift and nonsense mutations. The mutations identified in milder cases have not only provided insights into thalassemia but also genetic evidence which beautifully complements and reinforces biochemical studies which identified key *cis*-acting elements important for transcription, splicing and polyadenylation. Hence in the 1970s and 1980s β-globin mutations were identified in the CACC box at 80–90 base pairs (bp) upstream of the transcription start site (−80–90), in the TATA box at −30 bp, the 5′ untranslated region (UTR), an alternative splice site and in the polyA addition site (Chapter 1).

Another interesting feature of thalassemias with relevance to common disease is that specific variants have reached high frequency in different populations, either through genetic drift or more plausibly due to selection driven by conferring resistance to malaria (Weatherall, 2001).

Long distance regulators

Ultimately, for whole genome SNP association studies it will be necessary to include variants at 1–3 kb intervals across the whole genome. At present this is not possible so difficult choices have to be made. One of the current approaches is to choose SNPs mapping within genes themselves, either non-synonymous variants, or splice site variants. However, we must be aware that there is increasing evidence of long distance regulators of gene expression. The first example of this came again from studying β-thalassemia. It was shown that deletions terminating up to 30 kb upstream of the transcription start site removed a key regulator

of the β-globin cluster termed the locus control region (LCR) (Grosveld *et al.*, 1987). Over the past few years there have been several more examples of such, so-called, position effects on gene expression in human disease (Kleinjan and van Heyningen, 2005). These position effects involve chromosome translocations or deletions which do not extend into the coding region of a gene, but lead to reduction of expression. This occurs either through removing or separating positive regulatory elements from the gene or by bringing negative regulatory elements into contact with the gene. One notable example, aniridia or congenital lack of an iris, results from haploinsufficiency of the *PAX6* gene, a key regulator of eye development. Most cases of aniridia arise through nonsense, splicing or frame shift mutations. However at least a dozen documented cases have arisen through downstream translocations or deletions with proximal breakpoints up to 250 kb away from the transcription termination site. These chromosome rearrangements have been shown to remove key conserved DNA sequences which are essential for the expression of *PAX6* in the developing eye. Remarkably, some of these regulatory elements are embedded in the intron of a neighbouring gene. Similar cases are now being reported on a regular basis (Table 10.1 and Kleinjan and van Heyningen, 2005).

The remarkable case of preaxial polydactyly: gain of function mutations operating on the sonic hedgehog gene from a vast distance

Few gain of function regulatory mutations have been documented in mammals. One very dramatic example concerns preaxial polydactyly (PPD) in humans and mice. PPD is characterized by extra digits in the fore and hind limbs. Normally, sonic hedgehog (SHH), the key regulator of digit formation, is expressed only in a posterior domain of the developing limb. In PPD, there is ectopic expression of SHH in the anterior part of the limb. One mouse mutant with PPD, *Sasquatch*, resulted from the insertion of a transgene 1 Megabase upstream of the SHH promoter. It was shown that the transgene inserted into a *cis*-acting regulator normally required for expression of SHH in the posterior limb domain (Lettice *et al.*, 2002). This cis-acting element covers 400 bp and is highly conserved across all vertebrate species. Remarkably, humans and mice with PPD were shown to have point mutations located at various positions throughout this regulatory domain (Lettice *et al.*, 2003). These point mutations cause ectopic activation of SHH in the limb by an, as yet, unknown mechanism. This regulatory element, known as the SRE, is located in the intron of another gene, *LMBR1*. Moreover, there are several other genes mapping between the SRE and SHH, but these do not come under the influence of the SRE. So we must be aware of a situation where an SNP might map in the intron of one gene, whilst exerting its effects on another gene that may map hundreds or even thousands of kilobases away.

This type of extreme situation may be very rare, only occurring in a handful of developmental genes: it is too early to tell. The vast majority of documented disease-causing mutations do not affect regulatory elements, and if they do, they are often very close to the gene itself. On the other hand, there is a great deal of ascertainment bias in genetics. It is much easier to identify mutations in coding regions, so these will be identified first. Moreover, regulatory mutations may give a different phenotype than the coding mutations and this will take longer to sort out.

Another new mutation target: exonic splicing enhancers

There is a family of proteins called the SR proteins that regulates splicing by binding to specific sequences in RNA exons known as exonic splicing enhancers (ESEs). Over the past few years it has been shown that many disease-causing mutations may map within and disrupt these splicing enhancers, leading to mis-splicing. The first such reported cases were mutations in the breast cancer susceptibility gene, *BRCA1* (Liu *et al.*, 2001). The most

Table 10.1. Position-effect genes in human diseases

Gene	Gene function	Domains/motifs	Disease	Distance of furthest breakpoint[a] (kb)	3′ or 5′ Side
PAX6	TF	Paired box and homeodomain	Aniridia	125	3′
TWIST	TF	…	Saethre-Chotzen syndrome	260	3′
POU3F4	TF	POU homeodomain	X-linked deafness	900	5′
PITX2	TF	Homeodomain	Rieger syndrome	90	5′
GLI3	TF	Zinc finger	Greig cephalopolysyndactyly syndrome	10	3′
MAF	TF	bZIP	Cataract, ocular anterior segment dysgenesis, and coloboma	1000	5′
FOXC1	TF	Forkhead	Glaucoma/autosomal dominant iridogoniodysgenesis	25/1200	5′
FOXC2	TF	Forkhead	Lymphedema distichiasis	120	3′
FOXL2	TF	Forkhead	Blepharophimosis-ptosis-epicanthus inversus syndrome	170	5′
SOX9	TF	HMG box	Campomelic dysplasia	850	5′
SRY	TF	HMG box	Sex reversal	3	5′/3′
SIX3	TF	Homeodomain	Holoprosencephaly (HPE2)	<200	5′
SHH	Signaling	…	Holoprosencephaly (HPE3)	265	5′
SHH	Signaling	…	Preaxial polydactyly	1000	5′
SHFM1	TF	DLX5/DLX6?	Split-hand/split-foot malformation	450	5′/3′
FSHD	??	…	Facioscapulohumeral dystrophy	100	3
HBB	Oxygen carrier	Globin	γβ-Thalassemia	50	5′
HBA	Oxygen carrier	Globin	α-Thalassemia	18	3′
Hoxd complex	TF	Homeodomain	Mesomelic dysplasia and vertebral defects	60	3′
LCT	Enzyme	Lactase	Lactase persistence	15/20	5′

Note: – TF = transcription factor.

[a] In the case of 3′ breakpoints, the distance refers to the distance from the breakpoint to the 3′ end of the gene or complex.

(taken from Kleinjan and van Heyningen, 2005)

surprising implication is that apparently synonymous, silent variants may in actual fact have a pathological effect through causing aberrant splicing.

The genome sequence pinpoints novel potential regulatory variants

The genome sequence, particularly the comparative genome sequence, has revealed a number of new features including non-coding RNAs such as micro-RNAs (Grosshans and Slack, 2002) and deeply conserved DNA stretches several hundred base pairs in length, often mapping between genes (Bejerano *et al.*, 2004). There is now a concerted effort being put into dissecting the functional roles of these novel features. Moreover, the hunt is now on to see if variants in these sequences could lead to disease or are associated with disease susceptibility.

A plethora of functional mutations

Given all the above considerations, we now have to contemplate a wide range of potential mutational mechanisms for disrupting gene expression and protein function. The levels of RNA and the protein product may be influenced by mutations affecting a variety of regulatory steps – transcription, processing, RNA transport, stability and translation. Some of the mutations affecting transcription may act at a vast distance from the gene they influence. When it comes to missense mutations we may have thought conventionally of those that alter amino acids directly involved in a key function, such as catalytic activity or DNA binding. However, we are now aware that proteins can be modified in an increasing number of ways including phosphorylation, glycosylation methylation, acetylation, ubiquitination, sumolation, neddylation, etc. Each of these modifications can profoundly influence the function or localization of the protein within the cell. Hence, missense mutations affecting the target amino acids for these modifications could have severe functional consequences.

The types of functional variants affecting a particular gene will differ depending on the architecture of the gene and the nature of the protein it encodes. In some cases there might be selection at the population level for particular variants, in other cases not. All these factors will come into play when assessing what might constitute the key functional variant that contributes to disease susceptibility.

The study of monogenic subsets give mechanistic insights into common disease, leading to novel therapeutic approaches

Alzheimer's disease (AD) is a neurodegenerative condition always associated with pathological structures in the brain called amyloid plaques. One of the main constituents of these plaques is the 42 amino acid peptide called amyloid-β (Aβ) which is produced by cleavage from a neuronal transmembrane protein, the amyloid precursor protein (APP). There are a number of rare Mendelian subsets of AD, all characterized by early-onset, highly penetrant disease. Individuals affected with the monogenic forms of AD also exhibit plaques, as in the common form of AD. Prior to the genetic breakthroughs described below, there was much debate about whether the Aβ protein and the plaques were a cause or consequence of the disease. Now it seems there is little doubt that the production of excess Aβ and the resultant plaque formation cause disease by an, as yet, unknown mechanism (Roses, 1996). The three genes mutated in monogenic AD encode APP itself, and the presenilins 1 and 2. The APP mutations are missense mutations affecting residues close to the point of cleavage of Aβ. These mutations lead to an increased propensity for cleavage and production of Aβ. The presenilins are part of a secretase complex that cleaves APP to produce various peptides including Aβ itself.

Not only have these studies proved a causal link between Aβ production and AD but they have also revealed a novel biochemical pathway and paved the way for new therapeutic approaches (Tanzi and Bertram, 2005). The first therapeutic approach being pursued is to identify drugs which specifically inhibit the ability of the secretase complex to produce Aβ. The second approach is immunological and based on primary studies in mice. Transgenic mice that overexpress Aβ in the brain were produced. These mice developed amyloid plaques and had impaired congnitive ability, as deduced by their ability to negotiate mazes and other tasks. Vaccinating the mice early in life with Aβ prior to developing symptoms prevented the formation of plaques and cognitive impairment (Janus *et al.*, 2000). Aβ vaccination is currently in phase 1 clinical trials in patients with AD. As with the mice, it would no doubt be more efficacious to vaccinate prior to the onset of disease in individuals known to be at high risk of AD. It is of course possible at the moment to identify individuals at high risk of AD, those homozygous for the *ApoE ε4* allele. It appears that all individuals might succumb eventually to AD if they live sufficiently long, i.e. well over 100. The *ApoE ε4* allele accelerates the onset of disease. On the other hand, the rare *apoE ε2* allele appears to be protective, delaying the onset of disease. A study of 1030 people aged 71–100 years in the Framingham cohort showed that 55% of *apoE ε4/ε4* homozygotes developed AD by age 80, whereas 27% of *apoE ε3/ε4* heterozygotes developed AD by age 85 and only 9% of those without an *ε4* allele developed AD by age 85 years (Myers *et al.*, 1996). Hence the risk ratio for AD was 3.7 and 30 respectively, for *ε3/ε4* heterozygotes and *ε4* homozygotes compared with non-ε4 carriers. No disease was detected in *ApoE ε2* heterozygotes or homozygotes. *ApoE ε4* appears to directly modulate cerebral Aβ oligomerization and clearance, again pointing to a common pathway for disease (Tanzi and Bertram, 2005). Currently it would be surprising if anyone would choose to be tested for their *ApoE ε4* status as very little could be done to treat the condition. However, if new pharmacological or immunological therapies bear fruit, the situation would change dramatically and it might be wise to screen the population in middle-age for ApoE ε4 status.

Familial hypercholesterolemia and statins

Perhaps the most dramatic example of monogenic diseases leading to treatments for common disease concerns familial hypercholesterolemia (FH) and the development of statins (Goldstein *et al.*, 2001). FH is a relatively common dominant condition affecting 1:500 of the population. Individuals with FH have high levels of cholesterol at birth and will develop atherosclerosis and vascular disease in childhood. The risk of fatal heart disease before the age of 40 years is significantly higher than among the general population. Approximately one-third of patients with FH show no symptoms until sudden cardiac death.

Through the pioneering work of Brown and Goldstein, we know that FH is caused by loss of function mutations in the LDL receptor gene (*LDLR*). The LDL receptor removes cholesterol from the blood stream. This piece of genetic detective work proved that high cholesterol is the cause of coronary disease in these patients. Extrapolating from this monogenic condition, these findings provide incontestable support for the epidemiological studies showing that high cholesterol is associated with coronary artery disease in the general population.

Brown and Goldstein argued that if cholesterol levels could be lowered in children with FH, this might prevent disease onset. To do this they screened for drugs that would inhibit the enzyme, HMGCoA reductase, which synthesizes most of the circulating cholesterol in the liver. This screen led to the identification and production of statins which reduced cholesterol levels in patients with FH and essentially cured them. Now of course statins are used to lower cholesterol levels in people at increased risk of coronary disease in the general population. Hopefully this remarkable

success story is a harbinger of many more such developments in the future.

Familial hypercholesterolemia as a model for the way environmental factors working together with single gene mutations modify phenotype

In a remarkable study, Sijbrands *et al.* (2001) were able to analyse the mortality rates associated with FH in a large Netherlands pedigree over the course of 200 years. The starting point was three probands who had mean fasting levels of serum cholesterol of 10.2, 9.2 and 12.8 mmol/l, respectively. All carried a V408M mutation on an identical haplotype suggesting they were distantly related. All maternal and paternal ancestors of the three carriers were traced and one pair of ancestors connecting all three probands was identified. Secondly, the descendents of this pair were traced and all living descendents were screened for V408M. The main finding as shown in Figure 10.1 was that mortality rates for individuals with this mutation varied dramatically over two centuries. In fact, mortality prior to the beginning of the twentieth century did not differ significantly from standard mortality rates and most individuals with the mutation had a normal life span. It was only during the twentieth century that mortality rates of carriers increased dramatically over the standard rate. The increase in mortality in the twentieth century is likely to have been caused by environmental factors such as smoking and high fat diet working in conjunction with the LDL receptor mutation.

Monogenic genes as genetic risk factors in common disease

One other potential benefit deriving from the study of single gene disorders is that it provides candidate genetic risk factors for common disease. Given the chequered history of association studies, which have often been underpowered, it is difficult to tell whether most of the claimed associations will stand the test of time. However, there are

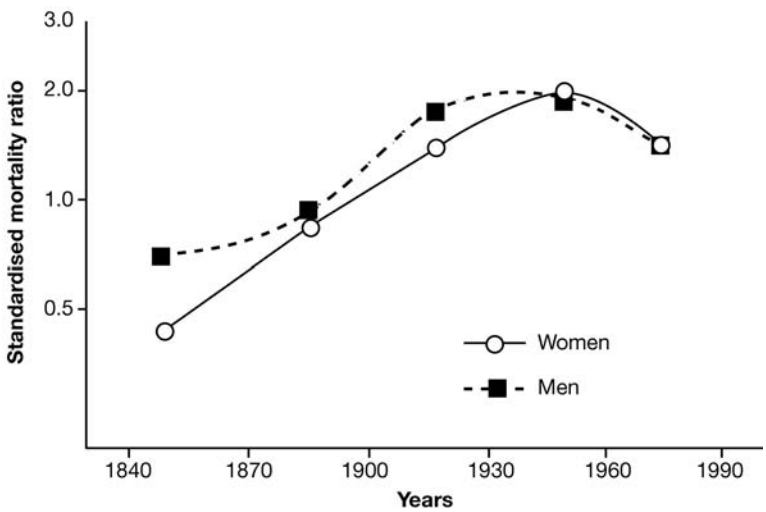

Figure 10.1 Mortality from familial hypercholesterolemia according to sex and time. Mortality was estimated among 250 persons with 0.5 probability of carrying the V408M. Probands and the first 20 years of life were ignored. From Sijbrands *et al.*, 2001.

a number of encouraging examples which have been replicated and where there is a convincing explanation for the mechanism by which the particular SNP exerts its effect on the disease risk.

Several such examples come from studies on the genetic basis of osteoporosis. Osteoporosis is a complex disease characterized by reduced bone mass (see Chapter 27), deterioration of the skeleton and increased risk of fracture. The most important clinical predictor of fracture risk is bone mineral density (BMD). Twin and family studies have shown that genetic factors account for 50–85% of the variation in BMD. One of the most convincing genetic risk factors for BMD is the type I collagen gene. Type I collagen is the most abundant protein in bone and is a heterotrimer made up of two collagen type I α1 protein chains (encoded by the *COL1A1* gene) and one collagen type 1 α2 chain (encoded by the *COL1A2* gene). Mutations in the coding region of *COL1A1* are found in 90% of cases of osteogenesis imperfecta, a rare condition associated with reduced BMD and multiple fractures. A polymorphism in the first intron of the *COL1A1* gene was found to be associated with BMD and restriction analysis has confirmed this result. Functional studies have shown that the risk allele increases binding to the transcription factor Sp1 leading to increased *COL1A1* transcription and an imbalance in collagen chains. Another interesting example concerns *LRP5*, an LDL receptor related protein which is a co-receptor for Wnt signalling. Mutations in *LRP5* are responsible for an autosomal recessive form of osteoporosis. Single nucleotide polymorphisms in the *LRP5* gene are associated with BMD in Japanese women (see Chapter 27). Other examples, where genes mutated in Mendelian subsets appear to be risk factors for common disease are quoted throughout this book. As only 1% of claimed associations in common disease can be replicated (L. Cardon, personal communication) genes mutated in monogenic subsets are likely to enrich for true candidates and simultaneously help to reveal mechanism.

REFERENCES

Bartek, J., Bartkova, J. and Lukas, J. (1996). The retinoblastoma protein pathway and the restriction point. *Curr Opin Cell Biol*, **8**, 805–14.

Bejerano, G., Pheasant, M., Makunin, I. *et al.* (2004). Ultraconserved elements in the human genome. *Science*, **304**, 1321–5.

Bates, G., Harper, P. and Jones, L. (2002). *Huntington's disease*. 3rd edn. Oxford: Oxford University Press.

Dalkilic, I. and Kunkel, L. M. (2003). Muscular dystrophies: genes to pathogenesis. *Curr Opin Genet Dev*, **3**, 231–8.

Farrington, S. M. and Dunlop, M. G. (2004). Familial colon cancer syndromes and their genetics. In Eeles, R. A., Easton, D. F., Ponder, B. A. J. and Eng, C. eds., *Genetic predisposition to cancer*. London: Arnold Publishers, pp. 315–31.

Fodde, R., Smits, R. and Clevers, H. (2001). APC, signal transduction and genetic instability in colorectal cancer. *Nat Rev Cancer*, **1**, 55–67.

Goldstein, J., Hobbs, H. and Brown, M. (2001). Familial hypercholesterolemia. In Scriver, C., Beaudet, A., Sly, W. and Valle, D., eds., *The Metabolic and molecular bases of inherited disease*. New York: McGraw-Hill. pp. 2863–913.

Grosshans, H. and Slack, F. J. (2002). Micro-RNAs: small is plentiful. *J Cell Biol*, **156**, 17–21.

Grosveld, F., van Assendelft, G., Greaves, D. and Kollias, G. (1987). Position-independent, high-level expression of the human β-globin gene in transgenic mice. *Cell*, **51**, 975–85.

International Human Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human genome. *Nature*, **409**, 860–921.

Janus, C., Pearson, J., McLaurin, J. *et al.* (2000). A beta peptide immunization reduces behavioural impairment and plaques in a model of Alzheimer's disease. *Nature*, **408**, 979–82.

Kleinjan, D. A. and van Heyningen, V. (2005). Long-range control of gene expression: emerging mechanisms and disruption in disease. *Am J Hum Genet*, **76**, 8–32.

Krawczak, M., Ball, E. V., Fenton, I. *et al.* (2000). Human gene mutation database – a biomedical information and research resource. *Hum Mutat*, **15**, 45–51.

Lettice, L. A., Heaney, S. J., Purdie, L. A. *et al.* (2002). Disruption of a long-range cis-acting regulator for SHH causes preaxial polydactyly. *Proc Natl Acad Sci USA*, **99**, 7548–53.

Lettice, L.A., Heaney, S.J., Purdie, L.A. *et al.* (2003). A long-range SHH enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum Mol Genet*, **12**, 1725–35.

Liu, H.-X, Cartegni, L., Zhang, M.Q. and Krainer, A.R. (2001). A mechanism for exon skipping caused by nonsense or missense mutations in BRCA1 and other genes. *Nat Genet*, **27**, 55–8.

Myers, R.H., Schaefer, E.J., Wilson, P.W. *et al.* (1996). Apolipoprotein E ε4 association with dementia in a population-based study: the Framingham Study. *Neurology*, **46**, 673–7.

McKusick, V.A. (1998). *Mendelian inheritance in man. A catalog of human genes and genetic disorders*, 12th edn. Baltimore, MD: Johns Hopkins University Press.

Pharoah, P.D., Antoniou, A., Bobrow, M. *et al.* (2002). Polygenic susceptibility to breast cancer and implications for prevention. *Nat Genet*, **31**, 33–6.

Roses, A.D. (1996). Apolipoprotein E alleles as risk factors in Alzheimer's disease. *Annu Rev Med*, **47**, 387–400.

Rozmahel, R., Wilschanski, M., Matin, A. *et al.* (1996). Modulation of disease severity in cystic fibrosis transmembrane conductance regulator deficient mice by a secondary genetic factor. *Nat Genet*, **12**, 280–87.

Sijbrands, E.J.G., Westendorp, R.G.J., Defesche, J.C. *et al.* (2001). Mortality over two centuries in large pedigree with familial hypercholesterolaemia: family tree mortality study. *BMJ*, **322**, 1019–22.

Tanzi, R.E. and Bertram, L. (2005). Twenty years of the Alzheimer's disease amyloid hypothesis: a genetic perspective. *Cell*, **120**, 545–55.

The Cystic Fibrosis Genotype–Phenotype Consortium (1993). Correlation between genotype and phenotype in patients with cystic fibrosis. *New Engl J Med*, **329**, 1308–13.

The US–Venezuela Collaborative Research Project and Wexler, N.S. (2004). Venezuelan kindreds reveal that genetic and environmental factors modulate Huntington's disease age of onset. *Proc Natl Acad Sci USA*, **101**, 3498–503.

Venter, J.C., Adams, M.D., Myers, E.W. *et al.* (2001). The sequence of the human genome. *Science*, **291**, 1304–35.

Weatherall, D.J. (2001). Phenotype-genotype relationships in monogenic disease: Lessons from the thalassaemias. *Nat Rev Genet*, **2**, 245–55.

Wright, A.F. and Hastie, N.D. (2001). Complex genetic diseases-controversy over the Croesus code. *Genome Biol*, **2**, 2007.1–2007.8.

Zielenski, J. and Tsui, L.-C (1995). Cystic fibrosis: genetypic and phenotypic variations. *Annu Rev Genet*, **29**, 777–807.

Zielenski, J., Corey, M., Rozmahel, R. *et al.* (1999). Detection of a cystic fibrosis modifier locus for meconium ileus on human chromosome 19q13. *Nat Genet*, **22**, 128–9.

11

Environment and disease

A. J. McMichael and K. B. G. Dear

Defining ''environment''

The ''environment'', considered in relation to health, typically refers to air pollutants, chemical residues in food, contaminated drinking water, radioactive wastes and so on. Conventionally, therefore, ''environmental health'' refers to research and policy in relation to the health risks posed by ambient environmental exposures. These exposures usually impinge on whole neighbourhoods, communities or populations and are therefore usually not under the control of individuals. This has important implications for prevention strategies.

Environmental risks to health from ambient physical, chemical and microbiological factors induce harm via direct physical, toxicological or microbial action. In addition to these various well-recognized human-made environmental exposures there are many important naturally occurring environmental health hazards. These include exposure to solar radiation, extreme weather conditions, and chemicals naturally present in drinking water (e.g. fluoride).

A more inclusive definition of ''environment'' includes the built environment and the social environment. Indeed, the influences of urban design, housing quality, material circumstances, social diets, socioeconomic conditions and social relations upon disease patterns have claimed increasing attention from researchers. Similarly, some commentators include variations in self-determined exposures (as through tobacco smoking, dietary choices and contraceptive hormone use) as part of the suite of environmental influences on health. This wider perspective becomes important, for example, in considering environmental influences on the rise of obesity in modern urban populations. This incoming tide of obesity is essentially an ''environmental system'' problem, reflecting fundamental, community-wide, changes in the ways of living of contemporary urban dwellers, resulting in disequilibrium between energy intake and energy output.

A general point to make early in this chapter is that much of the environmental influence on population health is of a pervasive kind. If air and drinking water are contaminated, then all members of the community are at risk from these exposures. Environmental conditions thus become background to the consideration of inter-individual differences in genetic risks to health. That is, the extent to which a particular genetic variation affects a particular health outcome is often dependent on the environmental potentiating or triggering of that genetic effect. A corollary of this is that joint genetic–environmental effects (discussed in more detail later in the chapter) vary over time and between places, as environments change. Indeed, human societies in general are currently in a phase of continuous and rapid environmental changes, even as their gene pools remain fairly constant. Hence, genetic alleles that may previously have had negligible impact on human biological functioning and health outcomes may, through environmentally enhanced expression, assume greater importance under new, unfamiliar,

environmental conditions. Genetic variants that predispose to obesity are an example. Hence, the estimation of the genetic contribution to some particular disease may vary hugely between populations and times, depending on the prevailing level of any such environmental potentiators.

Given this ongoing change in environmental conditions, it is instructive to review, briefly, the historical path via which environmental health hazards have evolved? In the western world, through to the eighteenth century, the pre-eminent longstanding environmental health hazards were malnutrition and famine, along with endemic childhood infectious diseases and recurrent epidemics. After the 1740s, in Europe, malnutrition and famines receded as the modern agricultural revolution began. The extreme urban crowding, insanitary conditions, and working-class poverty due to early industrialization in the nineteenth century resulted in infectious diseases becoming the pre-eminent environmental health hazard. Subsequently, with the rise of modern large-scale industry and of synthetic organic chemistry in the twentieth century, contamination of local air, water, soil and food became the major focus of environmental health concern. This happened, first, in developed countries and has been happening more recently in developing countries.

Today, as the scale of the human enterprise intensifies, we are encountering a new dimension of environmental health hazard, extending to regional and global levels (McMichael, 2002). This is exemplified by global climate change, stratospheric ozone depletion, and the disruption of food-producing ecosystems. Accordingly, environmental hazards to health are no longer mostly confined to the geographic location in which the initiating human action originally occurred.

Estimations of the health impact of environmental factors

There is abundant evidence that ambient environmental factors influence the incidence of many diseases. This is attested to by diverse epidemiological studies showing marked shifts in disease risks in migrant populations who transfer between environments – but whose genes are, of course, held constant. The World Health Organization (WHO) has recently reported a comprehensive assessment of global and regional burdens of disease and premature death, apportioning the burden between 26 major risk factors (World Health Organization, 2003). The set of environmental factors used by WHO includes unsafe drinking water, poor sanitation and hygiene, urban air pollution, indoor air quality, environmental lead exposure, and climate change. For the high-mortality members of the developing country category, an estimated 11% of total loss of disability-adjusted life years (DALYs) was attributed to those environmental exposures. For the low-mortality developing countries and for developed countries, the figures attributed to environmental factors were 6% and 2%, respectively.

The relative importance of environmental exposures as a cause of human disease and premature death remains contentious. Depending on definitions and assumptions, estimates of the environmental contribution to the global burden of disease and premature death vary. The World Health Organization has estimated that about 25% of the global burden, as measured in disability-adjusted life years (DALYs), is caused by environmental hazards, along with around one-sixth of the total burden in children (World Health Organization, 1997). Smith and colleagues (Smith *et al.*, 1999), in an analysis that encompassed disease initiation, progression, and case outcome, estimated that 25–33% of the global burden of disease and premature death is attributable to direct environmental risk factors.

Such estimates generally take account of only one kind of environmental health hazard, that arising from identifiable deleterious agents such as pollutants. The impact of such a hazard is relatively readily defined, even if not always easily measured: that is, we simply wish to compare the health impact of the actual, hazardous, environment with

the ''null'' environment in which the hazard is absent. However, there are two other types of environmental hazard that are not usually included in such estimation projects.

First, the environment may be deficient in something necessary to good health. Examples include dietary micronutrients, such as iodine or selenium. Second, a mismatch may have arisen between a particular local environment and the people now living in it. For example, migration may cause people with an ancestry long-adapted to high sunlight levels to suffer from vitamin D deficiency when they migrate to high latitudes where there is insufficient ultraviolet radiation for their phenotype. The impact of the environment in such cases is less easy to define, because it is not always clear what corresponds to zero exposure to the environmental hazard.

These details aside, this type of macroscopic estimation of disease burdens attributable to environmental factors can be done in relatively reductionist fashion at the population level, mostly without needing to consider any interplay with genetic factors. (In some special cases, the whole-of-population genetic profile may render a population more or less susceptible to environmental influences.) However, in general, the interest in genetic influences applies at a within-population level, and is particularly relevant to understanding the risks faced by individuals as they encounter their local daily environments.

Before exploring the nature of the health risks experienced by individuals, from genes and environment, it is relevant to stand back from detail and review the broad processes by which environmental conditions influence both genotype and risk of disease.

Environment as determinant of (i) genotype and (ii) disease

The environment assumes two important, different, roles in relation to genetic influences on disease.

First, within a long evolutionary sense, various regional differences in the biological attributes of human populations have arisen as a result of natural selection pressures exerted by local environments on human biology. In this sense, environment acts as a determinant of the local human genotype. Many examples are well known in the anthropological and biomedical literature. They include: (i) variation in skin pigmentation, which (especially in today's world in which there has been considerable redistribution of regional populations over recent centuries) contributes to variations in local population risks of skin cancer and vitamin D insufficiency; (ii) various enzymatic polymorphisms which are associated with altered risks of various diseases (e.g. the N-acetylation pathway enzymes, various of the cytochrome P450 oxidative enzymes, the activity of alcohol dehydrogenase, and lactase activity) many of which such polymorphisms are presumed to reflect the allele-selecting impacts of ancestral dietary differences in regional populations (McMichael, 2001); (iii) the various haemoglobinopathies (such as sickle cell anemia and thalassemia) and red blood cell metabolic abnormalities that have been positively selected via the heterozygote's reduced risk of parasitism by the malaria-causing plasmodium.

Thus, evolution permits genetic adaptations to arise which offset the dominant environmental threats to health in each location. This is the first, and most fundamental, way in which environmental variation impacts on genes and risks of disease.

Second, the environment presents many exposures that directly alter the probability of occurrence of various diseases. Ionizing radiation contributes to the risks of breast cancer and leukemia. Asbestos causes asbestosis and mesothelioma. Environmental tobacco smoke increases the risks of lung cancer and coronary heart disease. Heavy metals induce various forms of organ dysfunction, for example, environmental lead exposure in early childhood impairs neurocognitive development. Patterns of exposure to various antigens in early childhood influence the maturation path of the young immune system, and also

induce specific sensitizations; subsequent environmental exposures to aeroallergens induce asthmatic attack in susceptible individuals.

While the varying pervasiveness of these environmental exposures influences the rate of disease in communities, the risk (from those exposures) to each individual within those communities is often modulated by individual genotype, that is, by individual ''susceptibility''. For this reason, we read often of the interplay between genetic and environmental influences, of ''gene–environment interactions'' (see also the preceding chapter). This is the usual notion of susceptibility: that in a high-risk environment, susceptible individuals will suffer more disease than will non-susceptible individuals. But the converse is also true: susceptible individuals will *only* suffer more in a high-risk environment. That is, in the absence of a sufficient external exposure, genetically based susceptibility is of no health consequence. This is therefore a second way in which the environment impacts on how genes determine disease.

Sometimes the environment affects the risk of a ''genetic disease'' in both of these ways at once. The following example illustrates how environmentally influenced rates of disease in whole populations are sometimes modulated by the population's evolutionarily acquired genetic profile.

Non-insulin-dependent (Type II) diabetes mellitus

The incidence of non-insulin-dependent (i.e. Type II) diabetes mellitus (NIDDM) is increasing in adult urban populations around the world. The approximately 3% of adults currently affected by NIDDM will become an estimated 5% by 2025 (World Health Organization, 1998), as populations age and urbanize, and as obesity becomes more prevalent. NIDDM is a serious metabolic disorder which damages kidneys, heart, blood vessels and retina. The disorder results from ''insulin resistance'' – a reduction in the sensitivity of bodily tissues to insulin, the hormone which coordinates the metabolic production, storage and mobilization of the body's main fuels: glucose and free fatty acids.

The more than tenfold difference in NIDDM prevalence between Pima Indians (Arizona, USA) and Polynesians at the high extreme, South Asian and West Africans in the middle range, and European populations at the low extreme probably reflects, at least in part, population differences in genetic susceptibility. Indeed, the persistence of elevated rates of NIDDM in South Asian migrants several generations after migration to the UK corroborates the importance of genetic factors (McKeigue, 1997). Meanwhile, the approximately fivefold difference in NIDDM prevalence between rural and urban populations in Tamil Nadu, southern India, living a mere 20 miles apart, attests to the role of environmental factors (McKeigue, 1997).

"Thrifty genes", and selective insulin resistance

Almost a half-century ago, when obesity was much less prevalent than today, only a few populations (Pima Indians, Nauruans and other Polynesians, and Australian Aborigines) appeared to be particularly susceptible to diabetes. Further, the two main types of diabetes – adult-onset (NIDDM) and child-onset (insulin-dependent diabetes, IDDM: distinguished by insufficient pancreatic production of insulin) – had not then been differentiated.

In 1962, J.V. Neel, a population geneticist, postulated that diabetes in susceptible populations was due to their ''thrifty genotype'', acquired evolutionarily in response to precarious feast-or-famine regimes (Neel, 1962). Neel argued that natural selection had favored those individuals with a heightened insulin release, thereby best able to store any temporary excess of dietary energy (principally by storing glucose as fatty acids in adipose tissue and as glycogen in liver). This thrifty genotype, argued Neel, subsequently became ''detrimental'' in the modern-world circumstance of continuous dietary abundance,

in which the resultant obesity eventually induces pancreatic exhaustion, insulin insufficiency and, hence, diabetes.

Subsequently, medical science differentiated NIDDM from IDDM. Further, insulin's actions on various metabolic pathways were elucidated. So, in 1982, Neel revised his hypothesis, arguing that certain populations had higher levels of genetically based insulin resistance. This genotype, he argued, had been acquired in response to hunter-gatherer diets which were relatively lower in carbohydrate, but higher in fat and meat, than were the predominantly vegetarian diets of our Australopithecine ancestors (Neel, 1982; Neel *et al.*, 1998; Brand-Miller and Colagiuri, 1994).

Indeed, Neel and others postulated that hunter-gatherer populations developed a selective insulin resistance, focused specifically on their carbohydrate metabolism. This would have maximized glucose levels in blood as an immediate fuel, and as substrate for fetal brain growth (a critical focus for natural selection), while maximizing the synthesis of lipids from dietary fatty acids and amino acids, and their deposition in adipose tissue as stored energy (Brand-Miller and Colagiuri, 1994; McMichael, 2001). The evidence for selective insulin resistance was further corroborated in the 1990s (O'Dea, 1991, 1992; Reaven, 1998).

Why, then, are European populations at relatively lower risk of NIDDM? Today, the accrued epidemiological evidence indicates that most, not just a minority, of the world's populations are prone to NIDDM in the wake of increasing obesity (Stern, 1991). European and European-derived populations are the main exception. Is it plausible that the postulated ancestral selective insulin resistance could have been particularly selected against in the forebears of today's low-risk ''European'' populations?

Several commentators have proposed that the distinctly earlier move to glycemic carbohydrate-based, agrarian diets by proto-Europeans in the Middle East, several millennia before the equivalent move in some other centres around the world, was an important factor (Brand-Miller and Colagiuri, 1994; Friedman, 2003). Further, this differential in glycemic-diet exposure is likely to have been compounded by the unique inclusion of (high glycemic index) dairy foods in that early proto-European agrarian diet (Allen and Cheer, 1995). Selection pressures could, for example, have acted on the resultant heightened tendency to gestational diabetes, with its seriously adverse consequences for fetus and neonate (Diamond, 1992).

A full account of the natural selection pressures in relation to NIDDM remains elusive. However, it is near certain that critical variations in dietary circumstances, over time, have evolutionarily shaped the pattern of differences in metabolic handling of dietary substrates. Viewed within this frame, environment is a formative influence on the human genotype. The complex dependence of immune system development and functioning on vitamin D synthesis and action may be another such example. This is illustrated by the epidemiology of multiple sclerosis, with its evidence of both environmental and genetic influences on aetiology.

Multiple sclerosis

Multiple Sclerosis (MS) is a chronic T-cell mediated demyelinating autoimmune disease of the central nervous system. Disease onset usually occurs between 20 and 40 years of age, after which progression is variable, but generally slow. The geographic distribution of the disease has led to much debate over the relative importance of racial susceptibility (presumably genetic) and environment. On the one hand, there is a strong gradient towards higher prevalence at higher latitudes, both north and south of the equator; on the other, there are marked deviations from this pattern, including a high prevalence among Sardinians and Palestinians, and low prevalence among Chinese and Japanese, black Africans, Maoris and Amerindians (Rosati, 2001).

The genetic epidemiology of MS has been well investigated. Concordance is 25% between

monozygotic (identical) twins, but only 2–3% between dizygotic twins, in full siblings, or in parent–child relations, confirming the substantial genetically based variation in risk which is hinted at by the racial variation in MS prevalence. Little is known of specific genetic risk factors, but it is thought that the *HLA-DR2* allele doubles the risk, and HLA-DR15 reduces the age of onset (Masterman and Hillert, 2002). However, the fact that the monozygotic concordance rate is far below 100% indicates that non-genetic (environmental) factors are also important (Dessa Sadovnick, 2002).

Suspected environmental risk factors for MS include vitamin D insufficiency (with attendant reduction in the immune-suppressive effect of vitamin D), late- or post-childhood infection with Epstein–Barr virus (Martyn *et al.*, 1993; Levin *et al.*, 2003) and lack of ultraviolet radiation (UVR) exposure. While the skin-damaging and carcinogenic effects of high frequency UVR are well established, it is now emerging that UVR also has several beneficial effects. In addition to completing the synthesis of vitamin D, by hydroxylation within the blood vessels of the skin, there is accruing evidence, mostly from animal experiments, along with corroborating evidence from human studies, that UVR affects various aspects of local and systemic immune system activity (Ponsonby *et al.*, 2002). More specifically, insufficient UVR exposure may contribute to various immune-related disorders in humans, including autoimmune diseases such as MS.

McMichael and Hall (McMichael and Hall, 1997), noting the early emerging evidence, suggested that the strong positive latitudinal gradient in MS prevalence in Australia might be due to the associated negative gradient in UVR. Freedman *et al.* (Freedman *et al.*, 2000) found that people who died from MS in the United States tended to have both low residential and low occupational exposure to sunlight. Preliminary evidence from a recent case-control study in Tasmania indicates a strong inverse relationship between lifetime UVR exposure and risk of MS (van der Mei *et al.*, 2003). It has also been suggested that risk of MS may be increased prenatally by maternal vitamin D deficiency (McGrath, 2001).

Given this, it might be expected that in any particular location, dark-skinned people would have higher rates of MS than light-skinned people. For a given level of UVR exposure, individuals with darkly pigmented skin are indeed more prone to developing vitamin D deficiency (Holick, 1994), yet White populations generally have higher rates of MS. For example, in the USA, the MS prevalence in African-Americans is half that of white Americans (Hogencamp *et al.*, 1997). The reasons for this apparent anomaly are not clear. Perhaps people with darker skin have other, compensatory, immunological characteristics which counter the effect of vitamin D deficiency on immunologic regulation. If so, this would be another example, as for diabetes discussed earlier, of different genetically based mechanisms having evolved in different environments, and these genetic differences in turn moderating the influence of the environment on disease risk. It is also likely that in mixed-race societies there are relevant socioeconomic differences between racial groups, with the poorer group encountering a higher level of exposure to infectious agents, especially viruses such as the Epstein–Barr virus, in early life. Current evidence indicates that it is the deferred contact with certain critical viruses that initiates the autoimmune disease process of MS.

This section focuses on how the environment combines with genetics to determine disease rates. We should not forget that the environment (including the social environment) can also act alone, in ways that preclude any important contribution from genetic variation. An important topical example is the steep rise in recent decades in obesity in western populations.

Research on the ‘‘causes’’ of becoming overweight or obese is instructive. Cross-sectional individual-level studies, including twin studies, indicate that about half the observed inter-individual variation in relative weight (body mass index or BMI) is attributable to genetic factors: the other half is due to variation in individual

behaviors (eating and activity) which determine individual energy balance. But the recent marked rise in prevalence of obesity in urban populations has been much too rapid to be explained by any change in population gene frequencies (evolution), and so must be attributed solely to environmental and behavioral changes. The ''half is due to genes'' statistic tells us only that, within a population living in an otherwise unchanging environment and culture, genes play a significant role in determining which individuals get fat. That statistic can tell us nothing about the contribution of genes to the changing profile of the whole population over decadal time – and this must be essentially zero.

Let us elaborate this example a little further. If society were absolutely static, and if all individuals behaved identically, then genes would explain 100% of inter-individual variation. If society is static (the implied assumption underlying cross-sectional studies), but individuals behave differently (assuming that their behaviour is unrelated to their ''obesity genes''), then one can attempt to apportion causal influence between genes and environment, as above. However, if society is changing over medium-term time, such that the underlying way of life for everybody is changing, then there will be a temporal trend in the proportion of persons getting fat. Strictly, that trend is attributable 100% to the changes in environmental conditions and associated culture, and not at all to genes.

Joint effects of genes and environment

When gene expression depends on environmental potentiation, we conventionally refer to ''gene–environment interaction''. However, debate persists among biostatisticians, epidemiologists, biomedical scientists and others as to just what ''interaction'' means and whether one can recognize it, and quantify it, from empirical research data (Clayton and McKeigue, 2001). It is therefore simpler, and appropriate, to talk of ''joint effects''. These comprise situations in which, for an individual, the disease-inducing impact of one factor manifestly depends on the individual's status on the other factor. This, in its more extreme form, is obvious: (i) individuals who are genetically ''gluten intolerant'' will develop celiac disease if exposed to wheat flour in the diet, but not otherwise; (ii) individuals with the sickle-cell genotype, if infected with the malarial plasmodium by mosquito bite, are less likely to develop life-threatening falciparum malaria than those without the sickling allele.

Note also that the phrase ''joint effects'' (as also for ''interaction'') implies no primacy of action. Each factor can be seen to condition the impact of the other. Epidemiologists have, in recent years, often preferred to talk of ''effect modification'', but this requires specification of the effector and modifier.

Two further examples of joint gene-and-environment effects follow. These invoke, also, a broader notion of environment: (i) meat-eating is both part of a social environment, varying between cultures and socioeconomic classes, and an expression of individual preference, and (ii) atherosclerosis is a surrogate measure of, similarly, prevailing social dietary environment and individual dietary behavior.

Epidemiological evidence indicates that a high intake of meat may increase the risk of colon cancer. However, the relationship is somewhat inconstant across study populations, being most consistent for red meat and processed meat (Norat *et al.*, 2002). Meat, well cooked at high temperatures, has an elevated content of heterocyclic amines, including several compounds known to cause colon tumors in experimental animals.

A recent case-control study (Roberts-Thomson *et al.*, 1996) suggested that this risk increase is largely limited to ''fast acetylator'' individuals, who carry at least one copy of the *F1* allele of the *NAT2* gene, which codes for N-acetyltransferase. In such people, having a meat intake in the highest third of the population increases the risk an estimated 3.5 times compared with people in the lowest third. In ''slow acetylators'', who comprise approximately

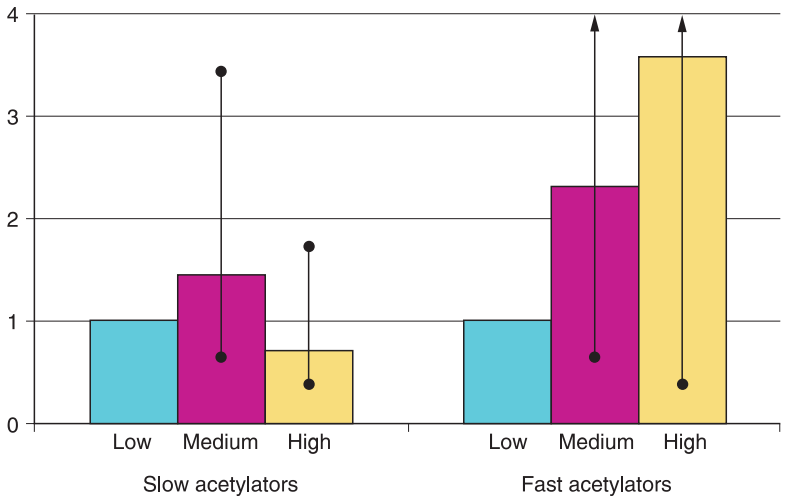

Figure 11.1 Risk of colon cancer or adenoma in relation to meat consumption and acetylator status. Case-control study, Adelaide, South Australia (from Roberts-Thomson *et al.*, 1996).

two-fifths of European-derived populations, there appears to be no increased risk at all (Figure 11.1). A plausible explanation for this pattern is that the heterocyclic amines in cooked meat are converted to mutagenic metabolites in the colon, and the concentration of these mutagens will reach dangerous levels only if: (a) the amines are present in abundance (due to high meat intake), and (b) the metabolic activation of the precursor compounds is efficient.

This particular joint gene–environment effect has not been found consistently in epidemiological studies. This could be for several reasons: (i) the N-acetyltransferase genotype may be confounded by other genetic polymorphisms which act in other directions, (ii) patterns of meat-eating differ between countries and cultures, (iii) the epidemiological classification of dietary habits, and especially attention to categories of meats eaten, is often inadequate. Nevertheless, the point of this example is to illustrate that, in some situations, there may be a clear cut differentiation of disease risk between those with and without a particular genetic allele. Such binary divides are the exception; most genetic polymorphisms confer a gradient in risk, including across molecular subtypes (defined by codon-specific location of mutation) of a particular ''allele''.

More commonly, we see a quantitative gradient in the environmental potentiation of effect across allelic subgroups of genotype. For example, the risk of developing Alzheimer's disease is raised by a factor of eight in individuals with an atherosclerosis score in the top quartile: but only if they carry at least one copy of the ε*4* allele of the *ApoE* gene (Hofman *et al.*, 1997). The risk in individuals with an atherosclerosis score in the top quartile and with non-ε4 alleles on both chromosomes is still increased, but only by a factor of three, relative to individuals in the lowest quartile of the atherosclerosis score (Figure 11.2).

Major Mendelian genes versus minor polygenes as predisposers

When ''genetic diseases'' are discussed, the usual focus is on the major single-gene disorders, such as cystic fibrosis, phenylketonuria and Huntington's chorea. In these disorders the affected individual is at very high, perhaps certain, probability of developing the disorder, no matter what. In such

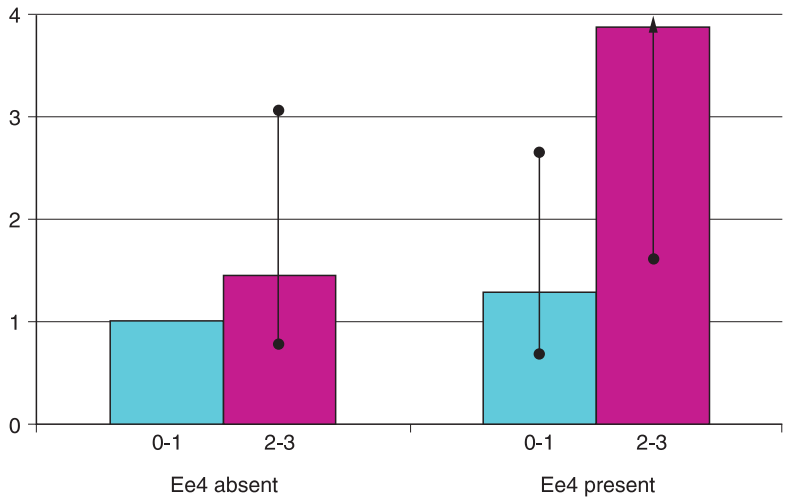

Figure 11.2 Joint risk of developing Alzheimer's disease as a function of apolipoprotein-E genotype and atherosclerosis score (from Hofman *et al.*, 1997).

cases, the interest in environment is either negligible, as an aetiological influence, or is special, as in the case of phenylketonuria where avoidance of phenylalanine in the diet can preclude occurrence of the metabolic disorder.

Most genetic influences on disease risks, however, occur in polygenic situations where any single contributory gene may shift the probability by no more than a few percentage points. Further, it is likely that it is the interacting *set* of alleles that is most relevant in determining the level of innate susceptibility of an individual to developing a disorder or disease in response to a particular environmental exposure. In recent years it has become apparent that there are multiple genetic loci which contribute to the occurrence of hypertension, a tendency to rapid weight gain, coronary heart disease, colon cancer, lung cancer and so on. This is hardly surprising. After all, the complex metabolism and physiology of the mammalian organism is, fundamentally, under genetic control, in that all proteins and other active molecules are genetically coded for, and the resultant slight inter-individual variations in the molecular structure of proteins affect their biological activity.

In light of this, it is important to realize that, for the great majority of individuals, any underlying genetic susceptibility to any particular disease is polygenic. ''Brave New World'' ideas of genetic ''correction'' in this context are premature, if not misplaced. The more important inference is that, by minimizing the potentially hazardous environmental exposure, the typically modest impact of those predisposing genes will be largely averted. That, of course, requires knowledge of the specific environmental triggers!

Prevention strategies: whole populations, high-risk groups or selected individuals?

There has been much recent discussion and debate over optimal strategies of disease prevention. On one axis there are important questions about the level, or stage, in the causal chain that is most propitious for effective intervention. Can we expect to stem the rise of obesity in modern urban populations by encouraging individuals to eat less and exercise more? Or should we be seeking to modify the living environment of whole

communities so that the opportunities and stimuli for eating energy-dense foods are lessened and the obligate amount of daily walking is increased?

On another axis, the crucial question, in relation to a population which has a manifestly elevated rate of some disease, is whether the focus of prevention should be on the identifiable minority of individuals within the population who are at high risk for that disease, or whether the whole population should be regarded, in the words of epidemiologist Geoffrey Rose, as ''sick'' (Rose, 1985). The latter approach leads to a more generalized intervention which seeks to shift the whole population risk distribution to a lower level.

That second issue goes further than that. As Rose eloquently pointed out, those factors that best explain the occurrence of cases within a population may not best account for the rate of the disease within the population at large. For example, if the population, overall, consumes a high-fat diet, then whether or not an individual smokes cigarettes may best explain whether he/she develops coronary heart disease. Meanwhile, the population-wide dietary behavior may be the main source of the elevated rate of coronary heart disease within that population. Consider another example, admittedly extreme, but it helps to make the point. If within a population everyone smoked 20 cigarettes per day, then the prime determinant of individual risk of lung cancer might well be one or more genetic polymorphisms which determine the fate of inhaled carcinogens. Yet it is the smoking that accounts for the overall high rate of lung cancer in that hapless population.

The extent to which a risk factor, whether environmental or genetic, is responsible for the observed variation in disease risks between individuals or subgroups is determined by two things: (a) the amount by which the factor increases the risk of disease in individuals subject to it, and (b) the extent to which the risk factor itself varies within the population under study. The research questions we can answer are thus constrained by the available variation in each factor.

Genetic studies on twins are valuable because they afford a degree of quasi-experimental control over the genetic variation: that is, we can ''hold constant'' the full genotype (monozygotic twins) or half the genotype (dizygotic twins). It is much less easy to achieve control over local environmental conditions. Nevertheless, care is needed in the interpretation of twin studies and other genetic-epidemiological studies. For example, a recent British study of twins reported that 68% of the inter-individual variance in childhood asthma occurrence was due to genetic factors, while only 13% was due to shared environmental factors (Koeppen-Schomerus *et al.*, 2000). The high proportion of risk here attributed to genes is misleading. First, whether the incidence of asthma within a given population is high or low, the extent of concordance between twins will be unchanged – and so the ''genetic'' component would be a constant. Yet any such variation in incidence depends on a non-constant factor: that is, variations in the constellation of environmental triggers for asthma. Indeed, if most or all individuals are exposed to an asthma-potentiating (modern, hygienic, urban-domestic) and asthma-triggering (air pollutants) environment, then the difference in actual asthma occurrence between individuals will substantially reflect genetic factors.

However, it is clear that genes cannot explain the marked rise in recorded incidence of childhood asthma that has occurred in developed societies over the past quarter-century. The much greater rates in today's society compared with those for people living a few decades ago reflects some fundamental changes in the environment. Viewed thus within that larger frame, the percentages attributable to genes and to the environment might well be reversed. The more important public health (and epidemiological) task is to explain those changes in disease rates at the population level. Within that particular context, the search for predisposing genes at the individual level is essentially irrelevant.

Conclusions

The onrush of modern molecular biology, and, in particular, molecular genetics, has raised hopes of pinpointing the genetic and resultant proteomic causes of human disorders and diseases. However, it is becoming apparent that there are relatively few one-on-one, single-gene, disorders. Most of the genetic susceptibility to disease within a population derives from the composite influences of a large number of genes that all bear on a particular aspect of metabolism or physiology.

Our genes provide us with a repertoire of capacities to respond to environmental exposures. Some of us over-respond; others under-respond. This, indeed, is inherent in nature: populations survive by dint of inter-individual genetic variability, such that there is always a minority of individuals able to survive an unexpected shift in environmental conditions. Thus does biological evolution proceed.

We should understand the perennial interplay between genes and environment in this light. Environmental factors act as triggers for many diseases, and they act more readily in the presence of multiple predisposing genes. The risks of those diseases can best be avoided by a combination of a community-wide reduction in exposure to that environmental factor (if that entails no other significant health penalty), and paying some special attention to the need to reduce the exposure of genetically higher-risk individuals. Such individuals can, now, often be identified by family history. In future, it is likely that we will be able to rapidly genotype individuals, and classify their risk profile in relation to a range of environmental factors.

At the population level, however, it is likely to remain the case that we will maximize the prevention of disease within the population by ameliorating the environment and other conditions of living, even as we become better able to fine-tune our efforts via knowledge of individual and family genotypes.

REFERENCES

Allen, J. S. and Cheer, S. M. (1995). ''Civilisation'' and the thrifty genotype. *Asia-Pacific J Clin Nutr*, **4**, 341–2.

Brand-Miller, J. and Colagiuri, S. (1994). The carnivore connection: dietary carbohydrate in the evolution of NIDDM. *Diabetologia*, **37**, 1280–6.

Clayton, D. and McKeigue, P. (2001). Epidemiological methods for studying genes and environmental factors in complex diseases. *Lancet*, **358**, 1356–60.

Dessa Sadovnick, A. (2002). The genetics of multiple sclerosis. *Clin Neurol Neurosurg*, **140**, 199–202.

Diamond, J. (1992). Human evolution; diabetes running wild. *Nature*, **357**, 362–3.

Freedman, D. M., Dosemeci, M. and Alavanja, M. C. R. (2000). Mortality from multiple sclerosis and exposure to residential and occupational solar radiation: a case-control study based on death certificates. *Occup Environ Med*, **57**, 418–21.

Friedman, J. (2003). A war on obesity, not the obese. *Science*, **299**, 856–8.

Hofman, A., Ott, A., Breteler, M. M. B. *et al.* (1997). Atherosclerosis, apolipoprotein E and the prevalence of dementia and Alzheimer's disease in the Rotterdam Study. *Lancet*, **349**, 151–4.

Hogencamp, W. E., Rodriguez, M., Weinshenker, B. G. (1997). The epidemiology of multiple sclerosis. *Mayo Clin Proc*, **72**, 871–8.

Holick, M. F. (1994). Vitamin D: New horizons for the 21st Century. *McCollum Award Lecture*, **60**, 619–30.

Koeppen-Schomerus, G., Stevenson, J. and Plomin, R. (2000). Genes and environment in asthma: a study of 4 year old twins. *Arch Dis Child*, **85**, 398–400.

Levin, L. I., Mungar, K. J., Rubertone, M. V. *et al.* (2003). Multiple sclerosis and Epstein-Barr virus. *JAMA*, **289**, 1533–6.

Martyn, C., Cruddas, M., Compston, D. A. (1993). Symptomatic Epstein-Barr virus infection and multiple sclerosis. *J Neurol Neurosurg Psychiatry*, **56**, 167–8.

Masterman, T. and Hillert, J. (2002). HLA-DR 15 and age of onset in multiple sclerosis. *European J Neurol*, **9**, 177–85.

McGrath, J. (2001). Does 'imprinting' with low prenatal Vitamin D contribute to the risk of various disorders? *Med Hypoth*, **56**, 367–71.

McKeigue, P. (1997). Diabetes and insulin action. In Kuh, D. and Ben-shlomo, Y. eds., *A life course approach to chronic disease epidemiology*. Oxford: Oxford University Press, pp. 78–100.

McMichael, A.J. and Hall, A.J. (1997). Does immunosuppressive ultraviolet radiation explain the latitude gradient for multiple sclerosis? *Epidemiology*, **8**, 642–5.

McMichael, A.J. (2002) Population, environment, disease and survival: past patterns, uncertain futures. *Lancet*, **359**, 1145–8.

McMichael, A.J. (2001). Diabetes, ancestral diets and dairy foods: An evolutionary perspective on population differences in susceptibility to diabetes. In Macbeth, H. and Shetty, P., eds., *Ethnicity and health*. London: Taylor & Francis, pp. 133–46.

Neel, J.V., Weder, A.B. and Julius, S. (1998). Type II diabetes, essential hypertension, and obesity as ''syndromes of impaired genetic homeostasis'': the ''thrifty genotype'' hypothesis enters the 21st century. *Perspect Biol Med*, **42**, 44–74.

Neel, J.V. (1962). Diabetes mellitus: a ''thrifty''genotype rendered detrimental by ''progress''. *American Journal of Human Genetics*, **14**, 353–62.

Neel, J.V. (1982). The thrifty genotype revisited. In Kobberling, J. and Tattersall, R., eds., *The Genetics of Diabetes Mellitus*, London: Academic Press, pp. 283–93.

Norat, T., Lukanova, A., Ferrari, P. and Riboli, E. (2002). Meat consumption and colorectal cancer risk: an estimate of attributable and preventable fractions. In Riboli, E. and Lambert, R., eds., *Nutrition and Lifestyle: Opportunities for Cancer Prevention*. Lyon: IARC (WHO), pp. 223–6.

O'Dea, K. (1992). Obesity and diabetes in ''The Land of Milk and Honey''. *Diabetes Metab Rev*, **8**, 373–88.

O'Dea, K. (1991). Westernisation, insulin resistance and diabetes in Australian Aborigines. *Med J Aust*, **155**, 258–64.

Ponsonby, A.-L., McMichael, A.J. and van der Mei, I. (2002). Ultraviolet radiation and autoimmune disease: insights from epidemiological research. *Toxicol*, **181–182**, 71–8.

Reaven, G.M. (1998). Hypothesis: muscle insulin resistance is the (''not so'') thrifty genotype. *Diabetologia*, **41**, 482–4.

Roberts-Thomson, I.C., Ryan, P., Khoo, K.K. *et al.* (1996). Diet, acetylator phenotype, and risk of colorectal neoplasia. *Lancet*, **347**, 1372–4.

Rosati, G. (2001). The prevalence of multiple sclerosis in the world: an update. *Neurol Sci*, **22**, 117–39.

Rose, G. (1985). Sick individuals and sick populations. *Int J Epidemiol*, **14**, 32–8.

Smith, K.R., Corvalan, C. and Kjellstrom, K. (1999). How much global ill-health is attributable to environmental factors? *Epidemiology*, **10**, 573–84.

Stern, M.P. (1991). Kelly West lecture: primary prevention of type II diabetes mellitus. *Diabetes Care*, **14**, 399–410.

der Mei, I., Ponsonby, A.-L., Dwyer, T. *et al.* (2003). Past sun exposure, skin phenotype and risk of multiple sclerosis: a case-control study. *BMJ*, **327**, 316.

World Health Organization. (1997). *Health and environment in sustainable development. WHO Report No. Document* WHO/EHG/97.8. Geneva: WHO.

World Health Organization. (1988). *1988 World health report*. Geneva: WHO.

World Health Organization. (2003). *World health report 2000*. Geneva: WHO.

12

Contemporary ethico-legal issues in genetics

Renate Gertz, Shawn Harmon and Geoffrey Pradella

Introduction

Since the inception of the current ''genetic age'' and its characterization as having created a ''risk society'', the ethical and legal debate surrounding the regulation of genomics (genetic research and related biotechnology and clinical genetics) has largely been informed by evaluations of risk. These risks pertain to individuals, populations, future generations and the environment. Cutting across these cohorts, the not unsubstantial risks of genetic relativism and determinism have excited concerns about genetic discrimination, triggered debates about the place and form of justice in the modern (genomic) health context, and challenged the practical realization and protection of human rights, individually, internationally and intergenerationally.

As evidenced by the papers which make up this volume, genomic knowledge is accumulating at an impressive, if not unparalleled rate. What is also evident, however, is that our comprehension of genomics remains somewhat rudimentary, particularly when it comes to understanding the interactions that occur between genes, and between genes and the environment. Further, while numerous genetic tests have been developed, we have been less successful at devising effective genetic therapies and treatments. That is not to say that genomics is a purely ''frontier'' science. On the contrary, it is becoming more and more a part of regular clinical practice, most particularly in the developed world. It has been claimed that there is:

> ... a transition [occurring] from traditional medical genetics, which focuses narrowly but effectively on heritable conditions, to genomic medicine, which integrates genetic information into everyday clinical practice. Indeed, according to proponents of genomic medicine, it is knowledge that will transform medical practice in the long-term – knowledge of how genes interact with each other and with the environment to cause disease (WHO, 2005).

If this is true, then it is vital that researchers and clinicians ('practitioners') have a level of ethical and legal awareness. This not only, and importantly, reduces the likelihood of injury to patients, but also safeguards the practitioner from liability. It is in that interest that this chapter addresses some of the ethical and legal issues, and regulatory responses, that have arisen in the following contexts: (1) genomics research (i.e. population and familial research); (2) genomics straddling the research/clinical settings (i.e. trials involving human subjects and the special considerations of pharmacogenetic research in the developing world); and (3) practical genomics in the clinical healthcare setting (i.e. the uses, risks and potential pitfalls of genomic applications).

Legal-ethical issues in research settings

With the mapping of the human genome, interest in population-based studies, such as biobanks,

has grown considerably within the medical community as these studies are part of the effort to bridge the gap between gene discovery, and the ability to utilize this knowledge for health improvement and disease prevention. The advent of these large-scale studies, however, has been accompanied by significant legal concerns which need to be addressed in tandem; issues such as privacy, confidentiality and consent. These legal concerns also vary for each type of population study. Family-based studies, for example, reveal different problems from general 'case-and-control' studies with randomly chosen controls. This section provides an overview of issues selected from amongst those suggested above, including recruitment issues for case-control and for family studies, as well as feedback for disease prevention, and more general issues regarding genetic data. Important legislation and regulation relevant to these studies includes the *Data Protection Act 1998*,[1] the *Human Tissue Act 2004*,[2] and the common law duty of confidentiality (within domestic United Kingdom law), and, on an international basis, the World Medical Association's Declaration of Helsinki (1964), and UNESCO's International Declaration on Human Genetic Data (2003).

Recruitment for case-control studies

As touched on in a later chapter (chapter 14), recruitment for 'cases' and 'controls' raises specific practical problems. Depending on the disease to be studied, disease registries, such as the cancer registries in the UK, are the most convenient source of recruitment for cases.[3] A list of eligible probands is established, and the people listed are approached, either via letter or a telephone call. This seemingly straightforward approach, however, brings with it certain legal and ethical problems. Data protection law will apply if the data held by the study can be considered ''personal data'', (i.e. if an individual can be identified);[4] and ''sensitive personal data'' if the data relates to the ''physical or mental health or condition'' of an identified or identifiable person.[5] As a result, the study will need to fulfil requirements set out in Schedules 2 and 3 of the Act, such as either informed consent from the participants, or that the research is in the public interest.

In order to contact potential participants, names and addresses need to be revealed. If the data is processed only for a single purpose (i.e. for one study, and not reused for other research projects), then the so-called research exemption of the *Data Protection Act 1998* applies. This principle states that research data must only be obtained for one or more specified and lawful purposes, and shall not be further processed in any manner incompatible with those purposes. The research exemption renders this further processing of data compatible with the purposes for which they had originally been obtained, and allows the data to be kept indefinitely. Two additional conditions also have to be met: the data are not to be processed to support measures or decisions relating to particular individuals; or, in such a way that substantial damage or substantial distress is, or is likely to be, caused to any data subject. If, however, the data are meant for several studies, data protection law requires either explicit consent of the data subject, or the fulfilment of one other condition, such as the protection of the 'vital interests' of the data subject, or another person, or the processing being necessary for the purpose of, or in connection with, any legal proceedings. Strict adherence to data protection law is of particular importance for obtaining and using data from Cancer Registries, as the resulting loss of trust could lead to a decline in patients agreeing to have their data entered.

[1] www.opsi.gov.uk/acts/acts1998/19980029.htm. (Accessed 7 February 2006)
[2] www.opsi.gov.uk/acts/acts2004/20040030.htm. (Accessed 7 February 2006)
[3] For an example, see: www.cancerindex.org/clink44x.htm. (Accessed 12 December 2005)
[4] Section 1(1) of the *Data Protection Act 1998*. See also, Information Commissioner (2002).
[5] Section 2(e) of the *Data Protection Act 1998*.

This, in turn, would lead to a decline in the importance of these registries.

Identifying and recruiting appropriate controls is a more difficult task than identifying and recruiting cases, as no single resource, such as a disease registry, can be exploited. As will be stated in Chapter 14, controls have to fulfil certain characteristics such as age, gender, ethnicity and residence. Researchers need to rely on primary care physicians and clinics to obtain the necessary information, or on volunteers responding to publicity. Frequently, for large-scale studies, an independent third party, (e.g. a team employed for this purpose) is used to generate a list of eligible participants, and to initiate contact. To comply with the legislation, the independent party regularly enters into an honorary contract with the United Kindom National Health Service (NHS).[6] This contract places the independent party on the same level as employees of the NHS, with the same obligations towards confidentiality.

Recruiting controls raises the same legal questions regarding data protection as does recruiting cases, and some additional problems too. A convenient method to recruit controls is the use of a sample cohort established for researching a different disease. For example, if the project intends to research breast cancer, controls could be recruited from a previously performed study on diabetes, where samples and data were kept for further use. If, however, these data are not completely anonymised, the problem of the secondary use of samples is raised, since the research exemption in the *Data Protection Act 1998* does not cover secondary use. Thus, using data for a purpose different from the one the data was initially collected for requires additional consent from the data subject (Lowrance, 2002). Obtaining informed consent from data subjects, perhaps several years after the samples and information were taken, may prove difficult. Guidance from professional health authorities and related councils suggests, as 'best practice', that whenever consent can be reasonably obtained, this should be done.[7] An example of 'unreasonable' in this context is the age of the data: if the data is several decades old, researchers cannot be expected to spend a considerable amount of time attempting to trace the participants, who may well have moved to a different address or even passed away. Another example is obtaining ethical approval of the study from a research ethics committee or through section 60 of the *Health and Social Care Act* 2001[8] and the Department of Health's Patient Information Advisory Group (PIAG).[9]

Informed consent is particularly difficult for genetic research databases, which are established as a resource for different research projects and contain identifiable data. An example of this is the UK Biobank.[10] Proper informed consent, in the truest sense, requires participants to be informed about each study for which their samples and data are used. If, however, at the time of sample and data collection, future research is still unknown, informed consent cannot be sought. Broad or open consent to genetic research, with the inclusion of a possibility for participants to withdraw, and the express exclusion of employers or insurance companies obtaining access to the samples and data, would then be one of the few feasible alternatives.[11]

Recruitment for family studies

Recruiting families for genetic studies brings to light additional problems to those encountered in recruiting participants for case-control studies. These additional problems also depend on the

[6] For examples, see: Department of Health (2003).

[7] For examples, see the guidance of the General Medical Council (GMC, 2002*a*), or the Medical Research Council (MRC, 2001).

[8] www.opsi.gov.uk/ACTS/acts2001/20010015.htm. (Accessed 7 March 2006)

[9] www.advisorybodies.doh.gov.uk/piag/About.htm. (Accessed 7 March 2006)

[10] www.ukbiobank.ac.uk/. (Accessed 7 March 2006)

[11] For a discussion of broad consent in genetic database research, see: Kaye (2004).

type of study, i.e. disease-specific or non-disease-specific. For both types, the recruitment process is divided into two parts: recruitment of the proband followed by recruitment of other family members. Likewise, for both study types, the recruitment methods for the proband are similar to those for general case-control studies; for disease-specific studies the methods for identifying and recruiting cases are likely to be applied, while for non-disease-specific studies, both controls and cases can be recruited.

Once the proband has been identified and contacted, two possibilities exist for recruiting other family members. Either the proband approaches family members, and discusses participation with them directly, or the proband provides researchers with their contact details (Beskow *et al.*, 2004). Both methods have ethical and legal implications. In the first scenario, the relatives are contacted by a person known to them, the proband, resulting in a likely reduction of privacy risks. In using this approach, however, the researcher is reliant on the proband to inform family members accurately about the study, and to contact family members, as promised. The question of whether a proband decides to participate, and returns the forms to researchers without familial consultations, or really has discussed the study with his or her family, remains problematic. A related problem is the possibility for coercion that such a scenario elicits, when the proband wishes to participate, but other family members do not (National Institutes of Health, OPRR, 1993).

In the second scenario, the proband provides researchers with contact details, handing over personal data to the researchers, such as family members' addresses and telephone numbers, and there is a risk that some family members consider this a breach of privacy (Botkin, 2001). Under the *Data Protection Act 1998*, as already mentioned, the processing of such personal data is only permitted when a condition in Schedule 2 is fulfilled, such as consent by the data subject, or where the processing is necessary for protecting the vital interests of the data subject.

A further problem relating to this type of study has to do with the number of family members that are recruited. If only part of the family is recruited, such as a sibling pair, the question arises as to whether other family members need to know, and, moreover, need to agree to the study, or the release of information. If the entire family is required, then it has to be determined whether consensus is needed, or whether the dissent of one family member can force an entire family to withdraw from participation. The answer to the second question will depend, largely, on the type of study to be undertaken.

For disease-specific studies, a different set of privacy considerations arises, for example the unavoidable or unintentional disclosure of a genetically linked disease to other family members. Several scenarios are possible. In one, the proband wishes to participate in the research project, but does not want other family members to know that he or she suffers from the disease. In this case, a choice has to be made between participating in the study, and not disclosing information. In a second scenario, the proband's instigation of discussions within the family may raise questions that expose a family member who, until then, had concealed having a particular disease. Perhaps the most difficult and serious issue is the possibility of disclosing a genetic disease to unsuspecting family members, who have not yet consented to participate, and who are therefore informed of something they may not wish to know, but which now can never be 'unknown' (Beskow, 2004). This last scenario speaks to the issue of the 'right not to know' unsolicited information which has the possibility of compromising psychological integrity.[12] Article 10.2 of the Convention on Human Rights and Biomedicine (Council of Europe, 1997*a*) states that ''Everyone is entitled to know any information collected about his or her health. *However, the wishes of individuals not to be so informed*

[12] For a detailed discussion, see: Laurie (2002).

shall be observed''.[13] The Explanatory Report to the Convention further justifies this 'right not to know' by declaring that ''patients may have their own reasons for not wishing to know about certain aspects of their health'' (Council of Europe, 1997*b*). When recruiting families for a disease-specific study, this ''right not to know'' will need to be adhered to. However, the practical application is fraught with difficulties, in particular due to the so-called public–private divide, where the law does not usually find applicability in the private core of a family unless ''invited in'' by a family member who feels that his or her rights have been breached. Researchers in this case will need to rely on the probands' knowledge of their family members to disclose or not to disclose facts which might prove to be traumatizing.

A unique problem raised by the need to recruit 'genetic families' arises where family members have been adopted, or are the offspring of gamete donation. Part of the solution lies in whether the children have full knowledge of the fact they are not genetically related to their parents, or are completely unaware. In the first case, no problem arises, but the second case generates the question of how the situation is addressed in the recruitment process without requiring, or creating the need for explanations or revelations that would, at least, be an unpleasant surprise, if not potentially harmful. A similar concern relates to a mother who has had extramarital relations, of which her husband and subsequently resulting child are unaware. If there was never any *medical* need for the child to be informed, then the result of disclosure in the course of a research study could be extremely traumatic. While adultery is certainly not legally protected, as is, for example, the anonymity of sperm donors, the effect of disclosure in such a case could be traumatic for the entire family (Gertz, 2004). This problem is similar to the 'right not to know', with researchers needing to rely on the probands and their families to avoid instigating a potentially traumatising situation.

[13] Emphasis placed by the author.

Feedback for disease prevention

Some studies, breast cancer studies in particular, offer participants an option to receive feedback from the study for possible disease prevention or early diagnosis purposes.[14] While an obvious benefit, it means that complete anonymization of the samples, and collected information, is impossible. The first issue with these studies, therefore, concerns privacy, and express informed consent will be required from all participants in order for the study to comply with data protection provisions. A problem with this approach, however, concerns siblings who have entered into such a study, with one sibling requesting feedback, while the other does not. In this case, it has to be ensured that the wishes of the sibling not requesting feedback are respected. By virtue, again, of the 'public–private divide', data protection legislation will not usually be applicable to interactions within a family, as the law is effective only in the public domain. The interaction between siblings touches the private domain, the inner core of people's families, and data protection law cannot be applied, therefore, unless 'invited in' by a family member who feels that their privacy has been breached to the extent that legal action is considered necessary. Problematically, researchers are thereby unable to ensure that the wishes of the sibling who does not want to receive feedback are adhered to within the family, that his or her right not to know is respected.[15]

A different legal and ethical problem arises in cases where the researcher is aware that the study participant has a sibling or children who might be affected by an inherited and heritable condition. The researcher may even feel compelled to warn the participant's family members about the possibility of their carrying the disease, possibly against the wishes of the participant. Parker and Lucassen give the example of a woman not wishing her pregnant sister to be informed about the woman's

[14] For an example, see: Richards (2003).

[15] For a detailed discussion of the right not to know, see: Laurie (2002), chapter 5.

son being diagnosed with Duchenne muscular dystrophy, as she believes this would lead to her sister terminating a pregnancy (Parker and Lucassen, 2003). Situations such as these place researchers in a considerable ethical dilemma, which can only be resolved on a case-by-case basis. Criteria to be weighed against each other include the obligation to respect the confidentiality of the initial research subject and the obligation to avoid harm and to provide information which may help the sister reach an informed decision. While the original data gained from the study refers to the participant herself, the problem lies in the fact that genetic data do not solely refer to a single person, but have implications for that person's blood relatives (Council of Europe, 2004). It could be said in this case that since these data also refer to the sibling, she had a claim to know the information (Mason and Laurie, 2006).

Ownership and governance issues

The advent of large-scale genetic studies has given rise to questions regarding the ownership and governance of blood samples, extracted DNA and related research results. In this regard, three stages can be identified where ownership and/or governance issues become a problem, namely: when participants donate their blood; when researchers analyse the DNA; and when research results become known. One of the most contested, and as yet unresolved, issues is the 'ownership' in blood and tissue samples. As the Medical Research Council guidelines state, the common law is unclear as to whether human biological material can be owned by anyone, and whether donors have any property rights over their samples (MRC, 2001). It is recommended, however, that researchers treat donated samples as gifts, in order to promote a 'gift relationship' between donor and researcher. While traditionally, altruistic reasons have been assumed for research participation, and commercial interests in the research outcome played no role (Laurie and Hunter, 2004; and Merz, 2002), this model of altruism is now being questioned. If a study results in profitable outcomes, such as patents, the feeling among participants may shift toward the expectation of a benefit (Weir, 2004; and Laurie and Hunter, 2004). This feeling can be exacerbated by misunderstandings about the nature of, and law governing property rights in samples. Many jurisdictions, such as the United Kingdom, do not recognize a property right in bodily material, yet evidence shows that participants frequently consider samples to be ''theirs'' in a property sense, and believe they have some legitimate claim to a return (Bovenberg, 2004).[16]

As soon as processing of samples and data provided by participants begins, the question of ownership arises a second time. Moreover, the issue grows in complexity, because in addition to the original data, or material, the data obtained through extraction and analysis of the DNA now needs to be taken into consideration. While, legally, it is fairly clear that the researchers will own any tissue or blood bank and will have other property rights, such as database rights, other models of governance, such as a type of custodianship (implying responsibility for safe storage of samples and safeguarding the donors' interests), have been proposed, and numerous examples are available through the appropriate research councils.[17]

The question of ownership in the results surfaces also if the research produces a patentable outcome, where the issue turns to the commercialization of the results, and the possible involvement of the pharmaceutical industry, and drug development. To be patentable, an invention must be new, must involve an 'inventive step' and must have utility or be made, or used, in any kind of industry. For several years, patents have been applied for, and granted, over human materials. Within the period of 1981–95 alone, over 1175 patents over human DNA were granted; and yet, the issue surrounding

[16] The tenor of such disputes is also well illustrated by the case of *Moore v. Regents of the University of California*, 793 P.2d 479 (Cal. 1990); *Cert denied* 499 U.S. 936 (1991).

[17] For examples, see: MRC (2001); and Nuffield Council (1995).

the patenting of life is still not satisfactorily resolved.[18]

Summary/conclusion

The discussion in this section of the chapter shows that studies drawing on the resources provided by genetic research databanks are accompanied by significant problems, beginning at the onset of the study, when participants are first recruited. Particular care is required with regard to data protection and privacy issues when families are recruited to take part in a genetic study, and feedback needs to be handled in a way that protects 'the right not to know', when and where appropriate. With regard to the ownership of donated samples, care needs to be taken to communicate a clear understanding of current or future property interests with participants, and to avoid confusion, samples should ideally be treated as gifts, donated by the participant. Finally, a statement regarding commercial interests should be added to the patient information leaflet.

Issues in clinical research settings

This section revisits some of the issues and questions just discussed, but in the context of human subjects research (HSR), which sharpens their import, because of the potential danger to the physical health of the participant. In addition, this section will touch on the special considerations which are relevant to multinational research, and research conducted in the developing world by entities from the developed world.

Research involving human subjects

HSR sits at the nexus of laboratory science and clinical practice. It is a necessary medico-scientific pursuit, in that it is key to translating the promise of biomedical research into improvements in clinical practice.[19] It is widely accepted, however, that HSR's propriety depends on its ethical foundation (i.e. the protection of the rights and well-being of participants and the existence of procedural safeguards related thereto). Unethical HSR, whether inadvertent, intentionally fraudulent or patently hurtful, can be damaging to individual participants, to public health more generally, and to the pursuit of knowledge through research.

Given this, many of the institutions which have an interest in HSR have drafted ethical codes that offer some guidance. Internationally, in addition to the United Nations Universal Declaration of Human Rights (1948), UNESCO's Universal Declaration on the Human Genome and Human Rights (1997),[20] and the Convention on Human Rights and Biomedicine (Council of Europe, 1997*a*),[21] there are a host of HSR-specific instruments:

- International Conference on Harmonisation (ICH) Tripartite Guideline for Good Clinical Practice E6(R1) (1996);
- World Medical Association Declaration of Helsinki (1964);[22]
- Council for International Organizations of Medical Sciences (CIOMS) International Ethical Guidelines for Biomedical Research Involving Human Subjects (2002); and
- Council of Europe Additional Protocol to the Convention on Human Rights and Biomedicine, Concerning Biomedical Research (2005).

Domestically, in the United Kingdom, HSR was long-governed by the common law of negligence and battery (liability for battery was avoided in the

[18] See, for example, the issues surrounding the patenting of the ONCO mouse in Canada, in Zahl (2004); and Kondro (2002).

[19] Indeed, Article 4 of the Declaration of Helsinki claims that medical progress is based on research which ultimately must rest in part on experimentation involving human subjects (WMA, 1964).

[20] UNESCO has also just completed its work on a more generalized declaration. See: UNESCO (2005).

[21] It should be noted that the UK is not a signatory to this convention.

[22] Arguably the most important statement of ethical practice in HSR.

medical context by securing consent), and by a complex matrix of non-statutory, governmental and professional guidelines which drew inspiration from the international instruments (Kong, 2004; and Ferguson, 2003).[23] Some of the relevant domestic guidelines for the UK are as follows:

- Royal College of Physicians (RCP) *Guidelines on the practice of ethics committees involved in medical research involving human subjects* (1997);[24]
- Medical Research Council (MRC) *Guidelines for good clinical practice in clinical trials* (1998);[25]
- Royal College of Paediatrics and Child Health (RCPCH) *Guidelines for the ethical conduct of medical research involving children* (2000);
- General Medical Council (GMC) *Good practice in research: the role and responsibilities of doctors* (2002*a*);
- Central Office for Research Ethics Committees (COREC) *Guidelines for researchers on patient information sheets and consent forms* (2005*a*);[26]
- Gene Therapy Advisory Committee (GTAC) *Operational procedures for the Gene Therapy Advisory Committee in its role as the National Ethics Committee for Gene Therapy Clinical Trials* (2005);[27]
- Department of Health (DoH) *Research governance framework for health and social care* (2005);
- MRC *Guidelines on good research practice* (2005*a*);[28]
- Wellcome Trust *guidelines on good research practice* (2005*a*).[29]

Although these guidelines are still operative (and must be complied with in specific settings), the governance situation has been clarified by the promulgation of the *Medicines for human use (Clinical trials) regulations 2004 (SI 2004, No. 1031)*.[30] Although the 2004 Regulations relate only to clinical trials of an investigational medical product (CTIMP), it is the stated policy of the UK Department of Health (DoH) that their terms are relied on in all types of research (DoH, 2005). Thus, the remainder of this section discusses some of the ethical considerations in HSR – ethical oversight, risk–benefit analyses, consent, remuneration, randomized trials, and publication – with reference to the 2004 Regulations.

Ethical oversight

HSR involves risks, and participants are asked to place a great deal of trust (and hope) in researchers and the research process. Responsibility for protecting participants is shared amongst the researcher, research staff, sponsors, contract research organizations and research ethics committees (RECs) (Sugarman, 2000).[31] In the UK, local RECs are established by individual Health Authorities to oversee projects involving participants from that authority. Multicentre RECs oversee projects involving participants from a number of authorities, or from abroad. Although

[23] A general lack of effective regulatory guidance and oversight with respect to various aspects of HSR has also been identified in other jurisdictions (Miller, 1998).

[24] The RCP has also published guidance on: *Research on healthy volunteers* (1988); and *Research involving patients* (RPC, 1990).

[25] See also: MRC (2005*a*).

[26] COREC operates in England only (with REC governance in Scotland, Wales and Northern Ireland falling to the relevant Ministries) and its work has recently been subsumed by the National Patient Safety Agency. See also: COREC (2005*b*).

[27] The GTAC website also provides a list of all gene therapy research up to 2003.

[28] For the MRC's most recent policy on scientific misconduct, see: MRC (1997).

[29] For the Wellcome Trust's most recent guidelines on handling research misconduct, see: Wellcome (2005*b*).

[30] www.opsi.gov.uk/si/si2004/20041031.htm (Accessed 7 March 2006). The 2004 Regulations, which came into force on May 1, 2004, are regulations under the *Medicines Act 1968*, and implement EU Directive 2001/20/EC. The MRC also offers advice for meeting the requirements established in the Regulations, and the EU has since supplemented the original Directive with Commission Directive 2005/28/EC, laying down further principles for good clinical practice as regards investigational medicinal products for human use.

[31] He concludes, after looking at the US situation, that institutions have an obligation to create an institutional culture that holds the protection of participants as paramount.

compositional and procedural matters are governed by DoH guidelines[32] and nationally applicable standard operating procedures (SOPs),[33] compliance with the 2004 Regulations is a practical necessity (Mason and Laurie, 2006, p. 659). Oversight must reflect the following:

- RECs must be comprised of people of both sexes, a range of ages and a broad range of expertise (including hospital staff, general practitioners, specialists and lay members to a maximum of 18 people (seven being a quorum) (Mason and Laurie, 2006).
- No researcher can commence a clinical trial unless a REC has given a favorable opinion.[34]
- RECs must consider a number of factors, including (1) the relevance of the clinical trial and its design; (2) whether the researcher's risk–benefit analysis is satisfactory and its conclusions justified; (3) the suitability of the protocol, researcher and support staff; (4) the quality of the facilities; (5) the adequacy, or completeness of the information supplied to, and the procedure of obtaining consent from, participants; (6) recruitment arrangements; (7) provisions for indemnity in case of injury or death; (8) researcher and participant remuneration agreements and sponsoring institution/facility agreements.[35]
- RECs can require researchers to resubmit their request if they are not satisfied with the information furnished.[36]
- Certain REC decisions can be appealed,[37] and are subject to judicial review.[38]
- The REC must be notified when (1) urgent safety measures are taken; (2) serious adverse events are experienced by participants; (3) unexpected serious adverse reactions are suspected; and (4) the trial is complete.[39]

Under the 2004 Regulations, it is an offence, punishable by fine and/or imprisonment, to (1) commence an unauthorized clinical trial; (2) fail to notify the REC and licensing authority of certain events (i.e. serious adverse effects); and (3) breach the conditions and principles of good clinical practice.[40] It is also an offence to provide RECs with false or misleading information, and the REC must report evidence of fraud or misconduct to its appointing authority (i.e. COREC, the MHRA).[41]

Risk–benefit analysis: Although protecting participants is a shared responsibility, the researcher is (and should be) the first and best guardian of rights and safety.[42] It is incumbent upon him or her to undertake a risk analysis, identify foreseeable risks and inconveniences (and weigh them against the benefits to participants and future patients), and structure the project in such a way as to minimize the risks to which participant(s) are exposed.[43] According to the RCP, patients should therefore only be involved in research that has greater than ''minimal'' risk where:

1. the risk is small in comparison to that already experienced by the patient as a consequence of the disease itself;
2. the disease is serious;
3. the knowledge gained from the research is likely to be of great practical benefit;

[32] For more on RECs, see: DoH (2001). For comment on the United States, see Burman *et al.* (2001).
[33] For more on SOPs, see COREC (2004); and White (2005*b*). Note that SOPs may vary throughout Europe. See: Hearnshaw (2004).
[34] Reg. 12. There are special requirements for obtaining authorization for trials involving medicinal products for gene therapy (Reg. 19) or with special characteristics (Reg. 20).
[35] Reg. 15(5). These are similar to the factors identified in COREC (2004).
[36] Reg. 15(2). Amendments and modifications to the authorization are dealt with in Regulations 22–26.
[37] Reg. 16.
[38] *R.* v. *Ethical Committee of St. Mary's Hospital (Manchester)* [1988] 1 F.L.R. 512.
[39] See: Regulations 30, 32 and 27 respectively.
[40] Reg. 49.
[41] Reg. 50.
[42] It has been argued that RECs and other institutional practices are of limited effect and that personal professional integrity is vital (Miller, 1998).
[43] See Reg. 15(5) and Schedule 1, Part 2, and Articles 16–18 of the Declaration of Helsinki (WMA, 1964). See also: Issa (2000), and Engelhardt (1996).

4. there is no other means of obtaining that knowledge; and
5. the patient gives fully informed consent. (RCP, 1990)

Healthy volunteers should only be involved if the project is of exceptional importance (benefit to society) and the research entails no more risk and/or burden for the participant(s) than is absolutely necessary. However, there is very little guidance as to how ''benefit to society'' should be assessed (Kong, 2004), and:

> There will clearly be differing views as to which risks are justifiable; it is equally clear that there will be those who would accept risks of a very high order ... and it is questionable whether they should be prevented from so doing. There are [however] legal limits to the extent to which consent decriminalises the infliction of harm... (Mason and Laurie, 2006)

Ultimately, the researcher must be alive to the risks, and must present those risks (and potential benefits) to the participant(s), whether they are patients or healthy volunteers, in a manner that is comprehensible.

Informed consent

The right to bodily integrity, which rests on the premise that everyone has the moral right of self-determination with respect to his or her body, has long been protected by the law. Unless consent to touching has been given, touching is a battery punishable both criminally and civilly (with an award of damages). Consent can make the touching lawful, but it is closely scrutinized by the courts (and other regulatory bodies), and this is just as true in the medical research context, where people (who are ill or may be in pain), are particularly vulnerable.

Informed consent obviously is an absolutely essential element of HSR. For consent to be ''informed'', the participant must be advised of, and understand, the purpose of the project, the benefits to him/herself and to society, the risks involved, and the alternatives available (Mason and Laurie, 2006). The 2004 Regulations stipulate that any adult able to give informed consent must be:[44]

1. interviewed by the researcher and given the opportunity to understand the objectives, risks, and inconvenience of the project;
2. informed of the right to revoke consent and to withdraw at any time;
3. provided with a contact point for obtaining further information; and
4. asked to give explicit informed consent.[45]

The 2004 Regulations probably should have also contained explicit requirements for ensuring voluntariness in terminally or chronically ill patients, disclosure provisions regarding the researcher's institutional affiliations and potential conflicts of interest, and provision of an independent contact point. These factors can be ethically important (and some of them are included in other guidelines), and researchers are well advised to address them, despite their explicit absence in the Regulations.[46]

Additionally, researchers have the responsibility of satisfying themselves that participants are competent. Where participants have reduced capacity or are legally incompetent (as in the case of mentally handicapped adults and minors), the 2004 Regulations erect additional safeguards and conditions which must be met.[47] For example, with respect to minors, researchers must interview the person with parental responsibility and ensure that they understand the objectives, risks and inconvenience associated with the trial, are given a contact person and informed of the right to withdraw the minor at any time. They must also

[44] Schedule 1, Part 3, clauses 1–5. Articles 20–23 of the Declaration of Helsinki address consent requirements in an (arguably) more comprehensive and demanding fashion (WMA, 1964).
[45] With respect to this condition, see COREC (2005*a*).
[46] For a critique of the consent provisions in the 2004 Regulations, see: Kong (2004).
[47] Schedule 1, Parts 4 and 5. Reference should also be made to the *Mental Capacity Act 2005*. See also Article 24–26 of the Declaration of Helsinki (WMA, 1964).

convey this to the minor and, where he or she is capable of assessing information and forming an opinion, also obtain his or her consent. Similar dual information provision and consent is necessary in the case of incompetents, the information and consent of the subject (again) being dependent on levels of comprehension and capacity, which must be assessed.

These safeguards are grounded on a desire to protect patients from exploitation; they must not be 'instrumentalized' (viewed simply as a means to benefit society).[48] Ultimately, it must be remembered that informed consent is not simply a mechanism for the protection of the researcher; it should be a genuine measure of the participant's comprehension and agreement to proceed, and it must be free of coercion or duress.

Remuneration

This concern revolves around both researcher and participant remuneration, and the potential conflicts that arise from that. Many of the concerns over the former relate to corporate research relationships.[49] Private entities, driven by commercial interests and pressing the clinical research envelope for efficiency reasons, can compromise the integrity of the project by creating circumstances in which scientific integrity, and professional clinical judgment concerning participant welfare, become influenced by financial considerations (Rettig, 2000). Even when direct conflicts of interest are absent, their very possibility may erode public trust. Thus, where private funders are involved, financial arrangements (both researcher-funder and institution-funder) should be disclosed for REC assessment. This is desirable because it allows best practices to emerge, and broad principles of equity between physician/researcher remuneration to be formulated (Ferris and Naylor, 2004).

Concerns over participant remuneration also relate to the potentially coercive nature of payments (undue inducement), which could jeopardize legitimate consent. Given the requirements for informed consent (which necessitate an understanding of the purpose, risks, benefits, alternatives and requirements of the project), and the many reasons people may have for participating, it may be that payments are not a threat to ethical acceptability (Grady, 2001), but there is no consensus on this point. The RCP characterizes patient payments as 'generally undesirable', except in the case of long and tedious studies, in which case, payment should not relate to risk, or act as inducement, but should relate to compensation for time, expense and inconvenience (RCP, 1990). In any event, in both cases, the 2004 Regulations require remuneration disclosure to the REC.[50]

Randomized trials

Randomized control trials (RCTs), which are designed to measure whether a new drug or treatment, is better than the existing one or none at all, are ethically controversial because they sometimes affront the duty to promote and safeguard patients' health, particularly where the researcher is also a primary carer.[51] For example, RCTs involve:

- administering a relatively untried treatment which may do harm to one group;
- withholding a treatment which may be of considerable benefit from one group;

[48] For more on informed consent and incompetent participants, see Wendler (2000). For special concerns relating to genetic testing and screening and minors, see Mason and McLean (2005).

[49] For a detailed discussion of these, see: Thompson (1993), Korn (2002), Bodenheimer (2002), Lexchin (2003), and Bekelman (2003).

[50] Regulations 15(5)(k) and (l).

[51] For this duty, see Articles 2, 3, 5, and 10 of the Declaration of Helsinki (WMA, 1964).

- administering inert substances (placebos) when there may be some alternative to the experimental treatment available;[52]
- making a scientific judgment as to whether the project should be discontinued (because of adverse reactions, hindrance of recovery or proof of effect or ineffect) (Mason and Laurie, 2006, p. 660).[53]

Ultimately, there is no satisfactory, or clearly correct, answer to these controversies, and RCTs remain ubiquitous. Under the 2004 Regulations, however, researchers must obtain a favourable REC opinion (and a Clinical Trial Authorization) prior to commencing the project.[54] In addition, the REC must revisit its favorable opinion in light of sponsor and researchers' progress reports (COREC, 2004).

Publication

Because divergent views on many aspects of research exist between researchers, RECs and the broader public, there is support for an ethical imperative to publish all research findings (including negative findings), both as a monitor of the researcher's ethics and as an audit of the REC's conduct in approving the research.[55] It has been argued that:

> There are … at least two unfortunate outcomes of [HSR] which, *a priori*, is not going to be published. [F]irst … when there is no accountability or potential scrutiny at higher levels, the subjects themselves may not be aware of the risks sequentially revealed as the research progresses. Problems surface only when … bodies are investigating, retrospectively, research that has gone blatantly wrong. The second unfortunate legacy of selective non-publication … is that anything that is potentially deleterious to a company's image or profit is suppressed.… This … means that all the inconvenience and risk to human research subjects is wasted, from the point of view of wider knowledge and future research (Pearn, 1995).[56]

Generally, therefore, it is suggested that RECs insist that researchers submit their results for publication in peer-reviewed journals.

Research undertaken in the developing world

It is generally believed that genetic research has an important role to play in improving healthcare globally (WHO, 2002*a*). HSR, particularly pharmacogenetic research, is increasingly conducted in the developing world, usually by entities from the developed world. Such multinational HSR raises ethical concerns in addition to those addressed above.[57]

Ethical recruitment from developing countries is particularly difficult because of the circumstances in which potential participants often live. For example, participants may have low levels of education and very different culturally based ideas about what the researcher is doing.[58] Also, participants will often be poor and only have access to rudimentary health care.[59] The potential for HSR, which offers treatment not otherwise available, to be unduly coercive, is very real, and added protections are therefore necessary

[52] The controversy over whether researchers are ethically required to offer ''the best proven method'' (which negates placebos) or, alternatively, ''the highest attainable and sustainable method'' of treatment in clinical trials has raged for some time. It loomed large in the revisions of the Declaration of Helsinki (WMA, 1964) and the CIOMS Guidelines (CIOMS, 2002). With respect to the former, see Article 29 and its Clarification.

[53] See also: Avins (1998), and Verdu-Pascual and Castello-Ponce (2001).

[54] Reg. 14–15, and DoH (2005).

[55] The ethical obligations of authors and publishers are addressed in Article 27 of the Declaration of Helsinki (WMA, 1964).

[56] The challenge of weeding out unethical research has also been discussed (White, 2005*a*).

[57] For examples, see: Nuffield Council (2002), and Nuffield Council (2005).

[58] MRC-sponsored vaccine trials in the Gambia which involved obtaining blood samples from subjects were seen by local participants, who clearly did not understand the nature or purpose of the research, as ''blood stealing''. (Fairhead, 2005.)

[59] In many developing countries, the average annual per capita expenditure on healthcare is less than US$2 with consequent massive disparities in access to healthcare. (Benatar *et al.*, 2003). See also: Gwatkin (2000), and Benatar (2001).

and warranted.[60] In short, the demands for achieving informed consent must be carefully considered because the validity of it will be more elusive, even though it may not be more critically scrutinized.[61] HSR in isolated communities with strong ethnic links and/or cultural traditions raises further challenges, because they may be more vulnerable to exploitation and discrimination. Studies suggest that recruitment and retention is improved when the community is involved in the risk–benefit analysis, and such analyses may necessitate involvement, or emphasis of social units (families); culturally specific implications of research are considered, and may only be identifiable by community members; and community concerns can be incorporated into research review mechanisms, without necessarily giving the community veto powers (Foster *et al.*, 1999).

Another unique feature of the developing world is the all-too-frequent absence of research governance. Contrary to the experience in the UK, and developed countries more generally, regulatory and ethical instruments and infrastructures are often lacking, as is the availability of biotechnological support tools. Unless scientific, health service and governance capacity can be facilitated through investment and collaborations, the existing social and healthcare inequalities tend to become exacerbated. Thus, researchers have greater obligations of sensitivity and vigilance with respect to meeting ethical requirements, like informed consent and voluntariness, and the added ethical responsibility to 'capacity-build' as an integral component of the project.[62] For research originating in the UK, the 2004 Regulations stipulate that where HSR is to be conducted in another country, the licensing authority may require the sponsor to undertake to permit the foreign premises to be inspected for the purposes of ensuring that conditions and principles of good clinical practice are satisfied, or obtain an undertaking from the owner to the same effect.[63] In addition, the sponsor must ensure that all suspected unexpected serious adverse reactions occurring at the foreign site are entered into a European database established under Article 11 of the *Clinical Trials Directive 2001/20/EC*[64] as a repository of information on such events.

Summary/conclusion

The discussion in this section further demonstrates that researchers must take care around issues of informing participants, obtaining consent, and disclosing affiliations and conflicts. They must seek prior ethical review, respond to ethical monitoring, and should publish all results, whether good or bad. It also demonstrates that researchers who undertake HSR, particularly when they are also primary care physicians, are in a very difficult position *vis-à-vis* their duties to the health and well-being of their patients, which may well conflict with their duties of scientific inquiry. This conflict is exacerbated when the HSR is located in developing countries, where the project may represent the best, or only, medical care available. Here they must be concerned with issues of justice, solidarity and health equity.

Legal-ethical issues in clinincal settings

Democracy cannot succeed unless those who express their choice are prepared to choose wisely.

Franklin Delano Roosevelt (1882–1945), 32nd President of the United States of America

[60] As recognized by Article 28 of the Declaration of Helsinki (WMA, 1964).

[61] Article 8 of the Declaration of Helsinki notes the special protection needed in HSR involving ''economically and medically disadvantaged'' people (WMA, 1964).

[62] This duty is explicitly addressed in CIOMS Guideline 20 (CIOMS, 2002). It is not addressed in the Declaration of Helsinki, but Article 19 (subject population should benefit from HSR) and 30 (participants should have access to treatments post-study) clearly implicates the notion of helping/advancing the subject community (WMA, 1964).

[63] Reg. 21.

[64] www.wctn.org.uk/downloads/EU_Directive/Directive.pdf. (Accessed 7 March 2006)

Embryonic screening and selection

Routine pre-implantation genetic diagnosis and screening is just over the horizon as an element of clinical practice, along with the capacity for the identification and 'de-selection' of embryos with genes linked to an increasing range of diseases, disabilities and basic traits. As with most advances in science, technology and knowledge, and as represented by the other chapters in this volume, newly emerging capacities to test and select genes at the embryonic stage of development present enormous prospective positives. These include dramatic improvements in the diagnosis, treatment and prognosis for debilitating, painful or terminal conditions, and opportunities to improve health and human productivity.[65] The potential and power of genetics, however, also needs to be approached with caution. As never before, genetic science is unlocking the intricate secrets that underlie the very fundamentals of our existence, essentially bringing us to the verge of a new era of man-made 'creationism'. From the enormous advancements realised in the last two decades, and those currently being pursued through research, new and daunting choices are, and will become, rapidly available within a social, political, biomedical, scientific and legal context, concomitant with the temptations, pressures and potential for their misuse (Larson, 2002). This latter possibility casts a long shadow over the current discourse, raising concerns akin to those which arose at the height of the early twentieth century's, arguably parallel, era of societal confidence in the capacities and potentials of genetic science: a confidence which, in that historic instance, spawned 'eugenics', and Draconian, if well-intended but largely coercive, eugenic practices that ultimately brought the movement to a discredited end.[66]

Drawing from these experiences, perhaps the greatest challenge in the twenty-first century is to anticipate the issues and pitfalls that our new-found technological capacities present, and from there, structure a framework that enables both individuals and societies to maximize the ethical and positive potentials for human and social benefit, while at the same time avoiding the individual and social harms that the earlier, ill-founded eugenic movements failed to prevent. Achieving these joint goals will require an approach, and a set of principles, that not only support already well-established law and ethics, but which, given the enormity of the consequences for future generations, ensure that the choices we make as a society today will be judged by history to have been 'virtuous', and judged by future generations to have been just.

The issues arising from preimplantation genetic diagnosis (PGD)[67] take on particular significance and relevance, for example, in relation to the potential effects they have on human rights protections, and societal attitudes toward persons with disabilities. As much as medical and diagnostic advancements promise to improve, treat or eliminate such conditions, they also have the potential to influence attitudes and responses to those who opt *not* to engage in testing or treatment, and those who are consequently born with genetically linked conditions. We have to recognise the responsibility that society has

[65] ''Human germline engineering is a discipline of the future. Its objectives will be to reverse genetic defects or to enhance desirable human traits...'' (Hood, 2000). See also: Micklos and Carison (2000).

[66] Modern usage of the term ''eugenics'' in reference to the selective breeding of human beings (intended for the ''improvement'' of human genetic qualities), was first formulated by Sir Francis Galton in 1865. The term is also used in reference to social and political movements which adopted or endorsed ''eugenic'' practices, and is broadly applied to any actions whose goal is to ''improve'' the human gene pool.

[67] The terminology and abbreviation varies in the literature, sometimes appearing as 'PGD' and also as PGS (preimplantation genetic screening).

to protect the rights and dignity of those who are, and would be, directly affected by the choices and actions of others, and we must be careful of adopting an overly optimistic and ahistorical perspective. An examination of the history of eugenics in the United States and the United Kingdom, among others, demonstrates that abuses can, and did originate in otherwise enlightened, liberal democratic environments, and that frameworks entirely formulated on voluntary practices and principles can become dangerously, if unintentionally, coercive.[68]

Making wise choices, then, will require a full and objective appreciation of robust ethical considerations within the social, legal and political context of modern democracies, a philosophical and legally just and fair basis on which to found those choices, and a determination of where lines need to be drawn between acceptable and unacceptable points on the spectrum. These include whether testing for certain conditions and not others risks implicitly creating a sanctioned or 'official' list of 'undesirable' traits. Further, whether the availability of tests targeted at certain traits is, in and of itself, a coercive influence on parental choice.[69] In this respect, we need to be mindful that there is a potential to harm persons currently living with disabilities. In striving toward a legitimate, just and virtue-focused effort to provide our children with the best possible chances for success, we have to ask at what point we risk their commoditisation as the end-products of esthetic and subjective choices relating to non-disabling characteristics and qualities (Robertson, 1994), or worse, tacitly consent to the inclusion of discriminatory non-therapeutic factors such as gender or sexual preference.

[68] For a comprehensive account of the history of eugenics in the United States, see: Pernick (1996).

[69] For example, when parents are faced with the prospect of having to decide whether they should continue (or initiate) a pregnancy that risks the creation of a person with a disabling or undesired condition, and in that regard seek, or are given, genetic or medical counseling.

Risks of coercion in counseling

It is widely accepted, in principle, that coercion of individual reproductive behaviors and choices, by governments or institutions, would be morally objectionable. Reproductive autonomy, or the right to make reproductive choices free from interference, has developed in recent years into one of the fundamental freedoms recognized in international law, and by a wide range of moral and ethical traditions.[70] This point is well reflected in the literature, and is a standard article in practice guidelines and conventions adopted by medical governing bodies, professional associations and international organizations. Accordingly, the American Society of Human Genetics reaffirmed its commitment to the principle of reproductive freedom, and ''unequivocally declare[d] its opposition to coercion'' in 1998.[71] The practice guidelines and policies of the American Medical Association are even more specific, stating that in their role as counselors and advisers, doctors should not impose personal moral values, or substitute their own moral judgements for those of their patients, when helping them understand their options, or when seeking consent (AMA, 2003).

[70] Rights to reproductive autonomy, or self-determination are arguably protected as a consequence of the operation of several specific, legally established human rights, including rights to life, liberty and security of the person, as found in the United Nations Universal Declaration of Human Rights (1948), Articles 1 and 3, the UN International Covenant on Civil and Political Rights (1976), Articles 6 and 9, the UN Convention on the Rights of the Child (1990), Articles 6 and 37, the European Convention on Human Rights (Council of Europe, 1950), Articles 2 and 5, the American Convention on Human Rights (OAS, 1978), Articles 4 and 7, the African Charter on Human and People's Rights (OAU, 1986), Articles 4 and 6, the Cairo Programme of Action of the United Nations International Conference on Population and Development (United Nations, 1994), and the Beijing Declaration, and Platform for Action (United Nations, 1995). For a fuller discussion, see WHO (2000).

[71] ''Knowledge-based decisions made by individuals or couples to avoid the birth of a child with disease or disability, so long as they are not unduly influenced by coercive governmental, institutional or other policies, are acceptable.'' (ASHG, 1999)

This does not mean, however, that actual practice is necessarily compliant, or that there is unanimous opinion as to whether medical, and particularly genetic counselling can ever be non-directive (Clark, 1991). Mary Mahowald points out, in a recent conference paper, that obstetricians are trained to be directive in the counseling of patients, a serious issue given the potential discrepancy in the power relationship between doctor and patient, especially when the patient is pregnant, and dependent for her care (Mahowald, 2003). Personnel providing direct medical services play a critical role in prenatal counseling, but it would appear that many such personnel do not necessarily practice in a manner that encourages objectivity, especially when the fetus at issue is potentially affected by a disabling trait (Asch, 2003). If this results in decisions effectively being directed by third parties, or those seen to be in positions of power, then even in an indirect or implicit manner, the principle difference between eugenics and modern genetic screening and counseling collapses.[72]

There also needs to be a high level of appreciation for, and sensitivity toward, paternalistic traditions and attitudes which are still powerful influences in society, particularly those supporting a paradigm that medical science is value neutral and authoritative. The status and authority accorded to the 'institution' of medical science is reinforced by the assumption that, as knowledge increases, so too does the ability to identify and treat illnesses and defects successfully. In terms of genetic screening and diagnosis, this can position doctors in the minds of their patients as powerful determiners (McLean, 1999). At the same time, we have to recognize that the ability to identify the genetic coding indicative of the presence of a particular disease or disability, such as Huntington's Chorea or Alzheimer's, does not extend to the ability to estimate the likely severity of the condition, should it develop, or the age of onset (Larson, 2002, p. 929). This means that counseling patients on the question of whether or not an identifiable genetic characteristic constitutes a serious risk, or a sufficiently serious condition to warrant embryonic de-selection, for example, is currently unanswerable, and therefore ethically problematic. In addition, many of the conditions for which we *are* able to identify the genetic hallmarks can vary significantly in their effect on the quality of life of the individual in question. The fact that the genetic factors for a condition can be identified through genotyping does not yet indicate, with a sufficient degree of certainty, whether, or to what extent, the condition will manifest itself, suggesting that a cautious and 'conservative' ethical and legal approach to the licensing of genetic diagnostic and selection technologies is well-advised.[73]

Attitudes and approaches to disability

The core criticisms that disability rights advocates have against PGD are (1) that it reinforces the 'medical' model that disability itself is the source of discrimination, not society's attitudes towards it (the 'social' model of disability) (Larson, 2002, p. 921); (2) that we make erroneous assumptions about the ability of disabled persons to live independent, productive and

[72] The current regime of pre-implantation genetic diagnosis and selection (PGD), for example, is arguably eugenic in its general purpose, and it is likely that some degree of eugenic practice is unavoidable. How the technology is legally and ethically guided and regulated, however, will dictate to what extent ''coerced'' eugenic consequences and decisions are beneficent and ethically sustainable.

[73] Note that the *purposes* of testing, and the availability of treatment are crucial factors differentiating the ethical and legal issues raised in this context. Testing directed at making therapeutic abortion or embryonic selection decisions, for example, are arguably ethically and legally separable from testing to assess treatment and assistance options, or to assist in parental preparedness.

fulfilling lives;[74] and (3) that we automatically label any lack of capacity or ability as 'bad' (Crossley and Shepherd, 2003).[75] Part of the problem with the search for a balance between realizing the benefits that genetic technologies present, and maintaining or securing legal and societal respect for disabled persons' interests and rights is, therefore, founded on the question of whether it is possible to reconcile principles of social inclusion while at the same time engaging in practices directed at reducing the numbers of certain persons in the community, specifically those with disabilities (Asch, 2003, p. 315). The extremely sensitive issue of whether it is 'better' to live without a disability is similarly at the core of the debate, and elicits serious concerns about the dominant construction of disability as a defect or deficit (Wilson, 2001). It also incorporates the characterization of genetic screening as something far removed from traditional medicine, in that it seeks to prevent *people*, rather than seeking to prevent or treat disease; that it replaces 'defective' with 'non-defective' individuals instead of seeking to provide therapy (Asch, 2003).[76]

Alternatively, if the ability exists to provide the best possible chances for one's children, and avoid suffering or lack of capacity, then why should parents not be allowed to make such choices? Some argue that genetic technologies should facilitate greater control over providing the best possible health for future children, enabling them to live out a typical lifespan, and equipping them with the best possible advantages so that they are able to realize all of the opportunities available to them. Is it possible, however, to suggest that having a disabling condition, limited basic human functions, or the risk of a reduction in the quality of one's health is 'undesirable', without stepping over a substantive ethical line? In a recent book on this subject, four bioethicists argue that, in a plural democratic society, it may not be possible, but that it is possibly unnecessary to set out a singular definition of what constitutes a good life, and that society should therefore allow for, and facilitate, the realisation of any number of divergent life choices (Buchanan *et al.*, 2000, p. 156). Referencing the eugenics period, they acknowledge the disastrous consequences, but suggest that society's current challenge is to set out to accomplish what the eugenicists of the twentieth century failed to do: a social, legal and ethical framework that provides for the optimal realization of genetic and human health and enhancement, mindful of successfully achieving the requirements of justice (Buchanan *et al.*, 2000, pp. 55–6).

On the opposite side of the issue, disability rights advocates point out that even strong legislative measures, such as the *Americans with disabilities Act 1990* (ADA)[77] in the United States, and the *Human rights Act 1998* (HRA)[78] in the United Kingdom, have failed to change the way science, medicine and bioethics approach disability questions, especially in the context of childbearing decisions (Asch, 2003, p. 329). On the contrary, they perpetuate the message of social stigma, that disability is abnormal, pathological, in need of correction, unfortunate, undesirable and to be avoided wherever possible. It is this form of 'better-off-dead' logic that disabled communities find so offensive, especially when the reality is that they are capable of experiencing a high quality of life, and rarely, if ever, constitute a public burden (Wilson, 2001, p. 176). There is also, they point out, one important aspect about which there has been little or no change over the last one hundred years, namely a real sensitivity to people with disabilities.

[74] Adrienne Asch argues that, ''These assumptions are demeaning to people with disabilities because they exaggerate their hardships and deprivation; obscure ... the discrimination they face; and ... discount their own testimony of living rich and rewarding lives.'' (Asch, 2003, p. 327.)

[75] See also: Asch (2003) pp. 317–18.

[76] See also: WHO (2002*b*), p. 165.

[77] www.eeoc.gov/policy/ada.html. (Accessed 7 March 2006)

[78] www.opsi.gov.uk/ACTS/acts1998/19980042.htm. (Accessed 7 March 2006)

Legislative protections aside, disability remains anathema to society in general, and efforts to eliminate it are encouraged. Indeed, the history of eugenics *is* the history of eliminating disability.

Equality of access and discrimination

In his recent book entitled *Our posthuman future*, Francis Fukuyama puts forward the fairly uncontroversial position that the basis upon which groups of people have historically been denied access, dignity and equal rights is through discrimination and prejudice (Fukuyama, 2002). In the nineteenth and early twentieth centuries, this manifested itself in assumptions that women were less rational, emotional, and therefore unsuited to professions or politics, and that immigrants from eastern and central Europe were less intelligent.[79] The eugenics movement at that time essentially relied upon the institutionalization and 'scientific legitimization' of such discriminatory attitudes in its programmes and proposals for sterilization, segregation and elimination of those with hereditary 'defects'. The risk we face with the advance of new genetic diagnostic technologies is that our improved ability to identify 'defects' will lead to increasing proportions of the population being labeled and treated as disabled or damaged, and therefore declared, at an embryonic stage, as worth avoiding. As the ability to discriminate, based on a host of genetic characteristics, broadens with the refinement of the technology, so, potentially, will the number of current and 'prospective' people susceptible to discriminatory practices and attitudes (Larson, 2002, pp. 923–33).

Perhaps a second, pertinent issue to be addressed, in order to (at least) avoid the creation of new societal divisions, is the danger of developing narrower genetic variations within distinct socioeconomic groups as a consequence of access to treatments and technologies being based on private financial resources (Fukuyama, 2002). Historically, there has existed a degree of socially and economically biased genetic selection through the choice of persons in marriage, divisions of space and association within communities, and the provision of broader and better opportunities to children from different backgrounds. In the future, however, limited access to genetic technologies and therapies presents the danger of embedding not only social advantages, but genetic ones as well, and the effective creation of a 'genetic upper class'. If genetic testing is not a generally accessible service, as part of a universal health programme, but is only available to those who can afford it, increasing inequalities will undoubtedly follow from the concentration of genetic 'problems' among the poorer members of the community (WHO, 2002*b*, p. 151). A strong argument in favour of universal access to basic healthcare being reinforced, and broadened to ensure principles of 'genetic equity and justice', is that, the 'genetic lottery' aside, inherited traits and genetic risks are simply morally undeserved.

Serious concerns have also been raised about the potential uses of PGD and genetic screening for *non*-therapeutic applications, like gender selection, and the possibility of its perpetuating or even amplifying gender discrimination. Accepting and endorsing the concept that perfectly healthy embryos would be discarded, or pregnancies terminated, merely because the fetus was the 'wrong' gender is prima facie unethical. Bioethicists, however, tend toward unanimous agreement that an argument exists in favour of gender selection when the motivation is the screening out of gender-linked genetic diseases (so-called 'therapeutic' selection). Some even propose that it might be an acceptable consideration for achieving gender-balanced family planning.[80]

[79] One of the bases on which the need for the American *Immigration Restriction Act 1924* was argued.

[80] The legal position in the United Kingdom on the use of PGD for family planning purposes, and in particular, the selection of a potential ''saviour sibling'' was determined by the House of Lords in the recent case of *R* (on the application of Quintavalle) *v. Human Fertilisation and Embryology Authority* [2005] UKHL 28.

Purely non-therapeutic motivations, however, are almost universally considered unethical, as evidenced in the strongly worded exclusionary references in medical guidelines and international protocols. The World Medical Association's *Declaration of Helsinki* states, in Article 23, for example, that medical research in tandem with medical care should only be undertaken if it is justified by potential diagnostic and therapeutic value (WMA, 1964). The Australian Medical Association's *Position Statement on Human Genetic Issues* (1998) is even more unequivocal, stating that any eugenic practices constitute violations of human rights, and that any genetic testing on preimplantation embryos should be limited to ''fatal or seriously and permanently disabling diseases.'' Further, that it is ''not ethical to practise genetic selection on the basis of gender'', open only to the exception of avoiding gender-related disease (Australian Medical Association, 1998).

Summary/conclusion

How PGD technology and its applications are guided and regulated in future will dictate to what extent their eugenic consequences are beneficent and ethically sustainable. Pressures influencing personal choice will inevitably arise from the relationship between public and private interests that will need to be accommodated, without their inappropriately influencing the direction of public policy and coercing the determination of individual, social and moral judgements. Belief in the principle that everyone in a free and democratic society deserves a fair and equal chance, and avoiding the risk of increasing existing social and economic inequalities, requires that societies demonstrate a commitment to these principles by ensuring that, as a minimum, PGD, other forms of genetic testing and screening, and genetic therapies and treatments are part of an accessible public healthcare system. While it is preferable for any system to err on the side of freedom, as opposed to government control or interference, especially when it comes to issues like personal autonomy, fundamental parental interests, and freedom of choice, also need to be balanced against the interests of the human lives, and the future generations that are ultimately affected and determined by the choices made. The legal status of unborn children and embryos, which are not currently recognized as 'legal persons', does not accord to them much protection. Given that modern genetic technology is uniquely and increasingly directive, if not determinative, of the very quality, personality and nature of those future lives, then perhaps the interests of unborn, and potential future children, in this context, are no longer quite so negligible, and so easy to legally ignore or subsume.

Conclusion

The consideration of legal and ethical issues in genetics and genomics serves to highlight a number of values which both the researcher and the clinician must bear in mind; notions of sanctity of life, human dignity, autonomy, justice and solidarity, for example, are implicated at every turn, and are emphasized to varying degrees in the different contexts considered. As further evidenced in this, unfortunately brief, survey, there are a large number of issues that must be considered essential elements in the development of objectives for, and the structure and operation of, future genetics research, genetically related clinical practices, and a regulatory regime to guide the conduct of genetic research, and its clinical applications. If these processes are dealt with appropriately, society should be able to optimize the potential benefits that can accrue from existing and future research, whilst protecting itself from a repeat of the nadir of the eugenics period. The greatest concerns with eugenics consistently relate to the coercive or compulsory means that characterized many of its programs. Therefore, a forward-looking structure of principles and mechanisms for the regulation of new genetic technologies should go beyond a reliance on the mere absence of explicit

coercion, and look toward a system of endemic principles, wherein legally and ethically defensible and sustainable choices can be derived from our interest of obtaining a sense of 'intergenerational justice'. Finally, while social consensus is critical to this process, we must also be mindful that popular accord does not constitute the basis for ethical validity, and the fact that practices might be, or might become widely accepted, does not make them morally justified.

REFERENCES

American Medical Association (AMA). (2003). Guideline E-2.12, in the AMA's *PolicyFinder*. www.insp.mx/Portal/MSP/bases-eticas/pdf/Declaraciones-materia.pdf. (Accessed 20 January 2006)

American Society of Human Genetics (ASHG). (1999). Statement of the board of directors of the American Society of Human Genetics: eugenics and the misuse of genetic information to restrict reproductive freedom (As approved, October 1998). *Am J Hum Gen*, **64**, 335.

Asch, A. (2003). Disability Equality and Prenatal Testing: Contradictory or Compatible? *Florida State University Law Review*, **30**, 334.

Australian Medical Association. (1998). *AMA position statement: human genetic issues. (Revised 2000. Revised 2002)*. www.ama.com.au/web.nsf/doc/SHED-5G7DAB. (Accessed 7 February 2006)

Avins, A. (1998). Can unequal be more fair? Ethics, subject allocation and randomised clinical trials. *J Med Ethics*, **24,** 401.

Bekelman, J. *et al.* (2003). Scope and impact of financial conflicts of interest in biomedical research: a systematic review. *JAMA*, **289,** 454.

Benatar, S. (2001). Distributive justice and clinical trials in the third world. *Theo Med*, **22**, 169.

Benatar, S., Daar, A. and Singer, P. (2003). Global health ethics: the rationale for mutual caring. *Int Aff*, **79**, 107.

Beskow, L.M. *et al.* (2004). Ethical issues in identifying and recruiting participants for familial genetic research. *Am J Med Genet A*, **130**(4), 424.

Botkin, J.R. (2001). Protecting the privacy of family members in survey and pedigree research. *JAMA*, **285**, 207.

Bodenheimer, T. (2002). Uneasy alliance: clinical investigators and the pharmaceutical industry. *New Engl J Med*, **342**, 1539.

Bovenberg, J.A. (2004). Inalienably yours? The new case for an inalienable property right in human biological material: Empowerment of sample donors or a recipe for a tragic anti-commons? *SCRIPT-ed*, **1**, 4.

Buchanan, A., Brock, D., Daniels, N. and Wikler, D. (2000). *From chance to choice: genetics and justice*. Cambridge: Cambridge University Press, p. 156.

Burman, W. *et al.* (2001). Breaking the camel's back: multicenter clinical trials and local Institutional Review Boards. *Ann Internal Med*, **134**(2), 152.

Central Office for Research Ethics Committees (COREC). (2004). *Standard operating procedures for ethics committees in the UK*. www.corec.org.uk/applicants/help/docs/SOPs.doc. (Accessed 6 March 2006)

Central Office for Research Ethics Committees (COREC). (2005*a*). *Guidelines for researchers on patient information sheets and consent forms*. www.corec.org.uk/recs/guidance/docs/Guidance_on_Patient_Information _Sheets_and_Consent_Forms.doc. (Accessed 28 November 2005)

Central Office for Research Ethics Committees (COREC). (2005*b*). *Information paper on informed consent in clinical trials of investigational medicinal products*. www.corec.org.uk/recs/guidance/docs/Informed_Consent_in_CTIMPsv1.1.doc. (Accessed 28 November 2005)

Clark, A. (1991). Is non-directive genetic counselling possible? *Lancet*, **338**(8773), 998.

Council for International Organizations of Medical Sciences (CIOMS). (2002). *International ethical guidelines for biomedical research involving human subjects*. www.cioms.ch/frame_guidelines_nov_2002.htm. (Accessed 1 December 2005)

Council of Europe. (1950). *Convention for the protection of human rights and fundamental freedoms*. Rome: Author. www.conventions.coe.int/Treaty/en/Treaties/Html/005.htm. (Accessed 7 March 2006)

Council of Europe. (1997*a*). *Convention for the protection of human rights and dignity of the human being with regard to the application of biology and medicine: Convention on human rights and biomedicine*. Oviedo: Author. www.conventions.coe.int/treaty/en/treaties/html/164.htm. (Accessed 8 February 2006)

Council of Europe. (1997*b*). *Explanatory report to the Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the application of biology and medicine: Convention*

on human rights and biomedicine. (ETS No. 164). www.conventions.coe.int/Treaty/en/Reports/Html/164.htm. (Accessed 8 February 2006)

Council of Europe. (2005). *Additional protocol to the Convention on Human Rights and Biomedicine, concerning biomedical research*. www.conventions.coe.int/treaty/en/Treaties/Html/195.htm. (Accessed 1 December 2005)

Council of Europe, Article 29 Data Protection Working Party. (2004). *Working document on genetic data, 12178/03/EN*. www.europa.eu.int/comm/justice_home/fsj/privacy/docs/wpdocs/2004/wp91_en.pdf. (Accessed 6 March 2006)

Crossley, M. and Shepherd, L. (2003). Introduction, genes and disability: questions at the crossroads. *Florida State University Law Review*, **30**, xi.

Department of Health (DoH). (2001). *Governance arrangements for NHS Research Ethics Committees*. www.dh.gov.uk/assetRoot/04/05/86/09/04058609.pdf. (Accessed 6 March 2006)

Department of Health (DoH). (2003). *Guidance note on research governance and honorary contracts*. www.research.luht.scot.nhs.uk/ResearchGovernance/Honorarycontracts.guidance.pdf (Accessed 6 March 2006)

Department of Health (DoH). (2005). *Research governance framework for health and social care*. www.dh.gov.uk/PublicationsAndStatistics/Publications/PublicationsPolicyAndGuidance/PublicationsPolicyAndGuidanceArticle/fs/en?CONTENT_ID=4108962&chk=Wde1Tv. (Accessed 28 November 2005)

Engelhardt, H. (1996). *The foundations of bioethics*, 2nd edn. New York: Oxford University Press.

Fairhead, J. (2005). When global technoscience meets poverty: pharmaceutical research and an economy of blood in the Gambia. Paper presentation at University of Edinburgh, 16 November 2005.

Ferguson, P. (2003). Legal and ethical aspects of clinical trials: the views of researchers. *MLR*, **11**, 48.

Ferris, L. and Naylor, C. (2004). Physician remuneration in industry-sponsored clinical trials: the case for standardized clinical trial budgets. *CMAJ*, **171**(8), 883.

Foster, M. *et al*. (1999). The role of Community Review in evaluating the risks of human genetic variation research. *Am J Hum Gen*, **64**, 1719.

Fukuyama, F. (2002). Unnatural selection: edited extracts from *Our Posthuman Future*. *The Guardian*, 14 May 2002.

General Medical Council (GMC). (2002*a*). *Good practice in research: the role and responsibilities of doctors*. www.gmc-uk.org/guidance/library/research.asp. (Accessed 28 November 2005)

General Medical Council (GMC). (2002*b*). *Research: the role and responsibilities of doctors*. www.gmc-uk.org/guidance/library/research.asp. (Accessed 6 March 2006)

Gene Therapy Advisory Committee (GTAC). (2005). Operational procedures for the Gene Therapy Advisory Committee in its role as the National Ethics Committee for gene therapy clinical trials. www.advisorybodies.doh.gov.uk/genetics/gtac/publications.htm. (Accessed 28 November 2005).

Gertz, R. (2004). Is it 'me' or 'we'? Genetic relations and the meaning of 'personal data' under the Data Protection Directive. *European Journal of Health Law*, **11**, 3.

Grady, C. (2001). Money for research participation: does it jeopardize informed consent? *Am J Bioethics*, **1**, 40.

Gwatkin, D. (2000). Health inequalities and the health of the poor: what do we know? What can we do? *Bull WHO*, **78**, 3.

Hearnshaw, H. (2004). Comparison of requirements of Research Ethics Committees in 11 European countries for a non-invasive interventional study. *BMJ*, **328**, 140.

Hood, L. (2000). The Human Genome Project. In Stock, G. and Campbell, J. (eds.), *Engineering the human germline* : an exploration of the science and ethics of altering the genes we pass to our children. Oxford: Oxford University Press, p. 17.

Information Commissioner. (2002). *Use and disclosure of health data: guidance on the application of the Data Protection Act 1998*. www.informationcommissioner.gov.uk/eventual.aspx?id=438. (Accessed 12 December 2005)

International Conference on Harmonisation (ICH). (1996). *ICH harmonised tripartite guideline for good clinical practice E6(R1)*. www.ich.org/cache/compo/475-272-1.html#E6. (Accessed 25 November 2005)

Issa, A. (2000). Ethical considerations in clinical pharmacogenomics research. *TiPS*, **21**, 247.

Kaye, J. (2004). Broad consent – the only option for population genetic databases? In Árnason, G., Nordal, S. and Árnason, V. (eds.), *Blood and data: ethical, legal and social aspects of human genetic databases*. Reykjavik: University of Iceland Press.

Kondro, W. (2002). Patenting life: Canadian High Court rejects ONCOmouse. *Science*, **298**, 2112.

Kong, W. (2004). The regulation of gene therapy research in competent adult patients, today and tomorrow: implications of EU Directive 2001/20/EC. *MLR*, **12**, 164.

Korn, D. (2002). Conflicts of interest in biomedical research. *JAMA*, **284**, 2234.

Larson, E. J. (2002). The meaning of human gene testing for disability rights. *University of Cincinnati Law Review*, **70**, 913.

Laurie, G. (2002). *Genetic privacy*. Cambridge: Cambridge University Press.

Laurie, G. T. and Hunter, K. (2004). Benefit-sharing and public trust in genetic research. In Árnason, G., Nordal, S. and Árnason, V. (eds.), *Blood and data: ethical, legal and social aspects of genetic databases*. Reykjavik: University of Iceland Press.

Lexchin, J. *et al.* (2003). Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ*, **326**, 1167.

Lowrance, W. (2002). Learning from experience: privacy and the secondary use of data in health research. The Nuffield Trust, pp. 19–20.

Mahowald, M. (2003). Aren't we all eugenicists? Commentary on Paul Lombardo's ''Taking Eugenics Seriously.'' *Florida State University Law Review*, **30**, 233. www.law.fsu.edu/journals/lawreview/downloads/302/MAHOWALD.pdf. (Accessed 20 January 2006)

Mason, J. K. and Laurie, G. T. (2006). *Mason & McCall Smith's law and medical ethics*, 7th edn. Oxford: Oxford University Press.

Mason, J. K. and McLean, S. A. M. (2005). Our inheritance, our future: their rights? *Int J Child Rts*, **13,** 255.

McLean, A. (1999). *Old law, new medicine: medical ethics and human rights*. London: Pandora, p. 8.

Medical Research Council (MRC). (1997). *Policy and procedure for inquiring into allegation of scientific misconduct*. www.mrc.ac.uk/pdf-mis_con.pdf. (Accessed 1 December 2005)

Medical Research Council (MRC). (1998). *Guidelines for good clinical practice in clinical trials*. www.mrc.ac.uk/index/publications/publications-ethics_and_best_practice/publications-clinical_trials_guidelines.htm. (Accessed 28 November 2005).

Medical Research Council (MRC). (2001). *Human tissue and biological samples for use in research: operational and ethical guidelines*. www.mrc.ac.uk/pdf-tissue_guide_fin.pdf. (Accessed 8 February 2006)

Medical Research Council (MRC). (2005*a*). *Guidelines on good research practice*. www.mrc.ac.uk/pdf-good_research_practice.pdf. (Accessed 1 December 2005)

Medical Research Council (MRC). (2005*b*). *Position statement on research regulation and ethics*. www.mrc.ac.uk/pdf-mrc_statement_regulations_ethics_may_2005.pdf. (Accessed 6 March 2006)

Merz, J. *et al.* (2002). Protecting subjects' interests in genetics research. *American Journal of Human Genetics*, **70**, 965.

Micklos, D. and Carlson, E. (2000). Engineering American society: the lesson of eugenics. *Nature Reviews*, **1**, 153.

Miller, F. *et al.* (1998). Professional integrity in clinical research. *JAMA*, **280**(16), 1449.

National Institutes of Health, Office for Protection from Research Risks (OPRR). (1993). *Protecting human research subjects: Institutional Review Board guidebook*.

Nuffield Council. (1995). *Human tissue: ethical and legal issues*. www.nuffieldbioethics.org/go/ourwork/humantissue/introduction. (Accessed 6 March 2006)

Nuffield Council. (2002). *The ethics of research related to healthcare in developing countries*. London: Author. www.nuffieldbioethics.org/go/ourwork/developingcountries/publication_309.html. (Accessed 7 March 2006)

Nuffield Council. (2005). *The Ethics of research related to healthcare in developing countries – a follow-up discussion paper*. London: Author. www.nuffieldbioethics.org/go/ourwork/developingcountries/publication_169.html. (Accessed 7 March 2006)

Organization of African Unity (OAU). (1986). *African (Banju) Charter on human and people's rights*. www.africa-union.org/official_documents/Treaties_Conventions_Protocols/BanjulCharter.pdf. (Accessed 7 March 2006)

Organization of American States (OAS). (1978). *American Convention on Human Rights (OAS Treaty Series No. 36, 1144 UNTS 123)*. www.oas.org/juridico/english/Sigs/b-32.html. (Accessed 7 March 2006)

Parker, M. and Lucassen, A. (2003). Concern for families and individuals in clinical genetics. *J Med Ethics*, **29**, 71.

Pearn, J. (1995). Publication: an ethical imperative. *BMJ*, **310**, 1313.

Pernick, M. S. (1996). The black stork: eugenics and the death of 'defective' babies in American medicine and motion pictures since 1915. Oxford: Oxford University Press.

Rettig, R. (2000). The industrialization of clinical research. *Health Affairs*, **19**(2), 128.

Richards, M. P. M. *et al.* (2003). Issues of consent and feedback in a genetic epidemiological study of women with breast cancer. *J Med Ethics*, **29**, 93.

Robertson, J.A. (1994). *Children of choice: freedom and the new reproductive technologies*. Princeton: Princeton University Press, p. 150.

Royal College of Paediatrics and Child Health (RCPCH). (2000). Guidelines for the ethical conduct of medical research involving children. *Arch Dis Child*, **82,** 177.

Royal College of Physicians (RCP). (1990). *Research involving patients*. London: Royal College of Physicians.

Royal College of Physicians (RCP). (1997). *Guidelines on the practice of ethics committees involved in medical research involving human subjects*. 3rd edn. London: Royal College of Physicians.

Sugarman, J. (2000). The role of institutional support in protecting human research subjects. *Aca Med*, **75**, 687.

Thompson, D. (1993). Understanding financial conflicts of interest. *New Engl J Med*, **329,** 573.

United Nations. (1948). *Universal declaration of human rights*. www.un.org/Overview/rights.html. (Accessed 1 December 2005)

United Nations. (1976). *International covenant on civil and political rights*. www.unhchr.ch/html/menu3/b/a_ccpr.htm. (Accessed 7 March 2006)

United Nations. (1990). *Convention on the rights of the child*. www.unhchr.ch/html/menu3/b/k2crc.htm. (Accessed 7 March 2006)

United Nations. (1994). *Programme of action of the United Nations International Conference on population and development*. Cairo: UN. www.iisd.ca/Cairo/program/p01000.html. (Accessed 7 March 2006)

United Nations. (1995). *Beijing declaration and platform for action – Fourth World Conference on Women*. www.un.org/womenwatch/daw/beijing/platform/. (Accessed 7 March 2005) Accessed and March, 2006, p. 47. (Accessed 6 March 2006)

United Nations Educational, Scientific and Cultural Organisation (UNESCO). (2003). *International declaration on human genetic data*. www.portal.unesco.org. (Accessed 9 February 2006)

United Nations Educational, Scientific and Cultural Organisation (UNESCO). (2005). *Universal declaration on bioethics and human rights*. www.portal.unesco.org/shs/en/file_download.php/46133e1f4691e4c6e57566763d474a4dBioethicsDeclaration_EN.pdf. (Accessed 7 March 2006)

Verdu-Pascual, F. and Castello-Ponce, A. (2001). Randomised clinical trials: a source of ethical dilemmas. *J Med Ethics*, **27,** 177.

Weir, R.F. and Olick, R.S. (2004). *The stored tissue issue*. Oxford: Oxford University Press, p. 289

Wellcome Trust. (2005*a*). *Guidelines on good research practice*. www.wellcome.ac.uk/assets/wtd002754.pdf. (Accessed 1 December 2005).

Wellcome Trust. (2005*b*). *Statement on the handling of allegations of research misconduct*. www.wellcome.ac.uk/doc_WTD002756.html. (Accessed 7 March 2006)

Wendler, D. (2000). Informed consent, exploitation and whether it is possible to conduct human subjects research without either one. *Bioethics*, **14,** 310.

White, C. (2005*a*). Suspected research fraud: difficulties of getting at the truth. *BMJ*, **381,** 281.

White, C. (2005*b*). UK agency to combat research misconduct. *BMJ*, **330,** 616.

Wilson, J. (2001). Disability and the genome: resisting the standardized genomic text. *Disability Studies Quarterly*, **21**(3), 168.

World Health Organization (WHO). (2000). *Considerations for formulating reproductive health laws*. www.who.int/reproductive-health/publications/rhr_00_1/considerations_for_formulating_reproductive_health_laws.pdf. (Accessed 7 March 2006)

World Health Organization (WHO). (2002*a*). *Genomics and world health: Report of the Advisory Committee on Health Research – Report by the Secretariat*. Geneva: WHO. www.who.int/gb/ebwha/pdf_files/EB111/eeb11112.pdf. (Accessed 7 March 2006)

World Health Organization (WHO). (2002*b*). Ethical issues in genetic research, screening and testing with particular reference to developing countries. In *Genomics and World Health: Report of the Advisory Committee on Health Research – Report by the Secretariat*. Geneva: WHO. pp. 147–73. www3.who.int/whosis/genomics/pdf/genomics08.pdf.

World Health Organization (WHO). (2005). *Genetics, genomics and the patenting of DNA: review of potential implications for health in developing countries*. Geneva: WHO. www.who.int/entity/genomics/FullReport.pdf. (Accessed 4 February 2006)

World Medical Association (WMA). (1964). *Declaration of Helsinki*. As amended and clarified (2004). www.wma.net/e/policy/b3.htm. (Accessed 20 November 2005)

Zahl, A. (2004). Patenting of 'higher life forms' in Canada. *Biotechnology Law Report*, **23,** 5.

SECTION 2

Common medical disorders

13

Developmental disorders

Stephen P. Robertson and Andrew O. M. Wilkie

An understanding of the complex processes that underlie the transition from zygote to newborn infant remains one of the major unsolved challenges in human biology. Failure of key steps in early embryogenesis leads to arrested development and embryonic wastage in a substantial proportion of conceptions (Wilcox *et al.*, 1999). Interference with later developmental pathways which mediate the processes of morphogenesis and organogenesis can also lead to fetal demise but equally can produce a phenotypic effect evident at term. This chapter discusses, with selected examples, our current understanding of the influence that genetic and environmental factors have on these complex developmental processes in humans.

The medical significance of developmental disorders

Developmental disorders in humans are diverse in nature and individually relatively rare, but as a group constitute a ''common disease''. Improvements in their recognition and pathogenesis, both as isolated entities and as components of syndromes, have been greatly aided by advances in the clinical speciality of dysmorphology and the construction of clinical databases which catalogue rare associations of phenotypic features (Donnai and Read, 2003).

The overall birth prevalence of disorders which are primarily considered to be due to defective morphogenesis is estimated to be between 2 and 3% (Kalter and Warnaky, 1983). If malformations associated with still births and abnormalities which do not present a requirement for significant medical intervention are included in this estimate, the figure rises to ~ 5%.

A list of some of the more common developmental anomalies is given in Table 13.1, together with an estimate of their individual birth prevalence. However, considerable variation in the birth prevalence of some disorders with time and location immediately raises suspicions of etiological heterogeneity. Whereas the prevalence of congenital heart defects adjusted for severity and ascertainment bias, has remained relatively constant over time and shows little variation between continents and over seasons (Hoffman and Kaplan, 2002; Bosi *et al.*, 2003), the prevalence of neural tube defects has altered dramatically over a time course of decades and shows marked fluctuation depending on geographical location, ethnicity, socioeconomic status and perhaps seasonality (reviewed in Botto *et al.*, 1999).

The relative contribution of developmental disorders to morbidity and mortality has altered over time with improvements in medical care (Yang *et al.*, 1997). The contribution to perinatal mortality that these conditions bring in developed countries is around 30%, with increasing numbers surviving into the neonatal period where birth defects remain a leading cause of death in children less than 15 years of age (Table 13.2). The economic and human cost of congenital

Table 13.1. Birth prevalence of selected developmental disorders in humans

Organ system	Specific disorder	Birth prevalence (per 1000 live births)
Cardiovascular system	Septal heart defects	2.5–3.5
	Complex heart defects	0.6
Central nervous system	Neural tube defects	1–3
Craniofacial anomalies	Facial clefting	1–1.5
	Craniosynostosis	0.5
Genitourinary tract	Hypospadias	3
Gastrointestinal tract	Tracheo-esophageal fistula	0.25
	Pyloric stenosis	1–3
	Diaphragmatic hernia	0.4
	Hirschsprung disease	0.25
	Omphalocoele	0.25
Skeletal	Congenital hip dislocation	5
	Club foot	0.5–0.9

Table 13.2. Mortality due to birth defects by age group, United States, 1979–1992

Age (years)	<1	1–4	5–9	10–14	15–24	25–34	35–44	45–54	55–64	>65	Total
Fraction of all mortality due to birth defects (%)	55.3	6.6	2.1	1.7	4.1	3.9	3.5	3.9	5.4	13.6	100
Mortality due to birth defects as a fraction of total deaths (%)	23.9	15.5	8.3	5.6	1.8	1.2	0.9	0.6	0.3	0.2	0.9

After Yang *et al.*, 1997.

developmental disorders is difficult to quantify but certainly is both substantial and significant.

Broad etiological groups

Studies of morphogenesis and embryology in the human and its application to birth defects identify three broad etiological groups (Spranger *et al.*, 1982). The terms ''disruption'' (indicating an alteration of a structure after or during formation by a destructive process) and ''deformation'' (the alteration of structure by the influence of an external mechanical force) can be distinguished from ''malformations'', which identify defects due to an intrinsic abnormality within a developing tissue acting during its formation. All three categories can exhibit environmental and genetic components, but the concept has been useful in understanding the pathophysiology of phenotypes and their relationship to etiology. There remain some clinically defined conditions (such as the VATER association – vertebral and anorectal anomalies, tracheo–oesophageal fistula, renal/radial defects) whose etiological basis remains enigmatic but whose definition owes much to this early morphology-based approach to disease classification.

Only a minority of developmental disorders have an unequivocally identified genetic basis. Chromosomal disorders account for 8–12% and a recognizable monogenic condition accounts for

Table 13.3. Genes identified as mutated in syndrome congenital heart disease

Disease	Gene	OMIM number for gene
Atrial septal defects	*EVC, EVC2, GATA4*	22550, 607261, 600576
Septal heart defects	*CBP, CRELD1*	600140, 607170
Atrial septal defect/tetralogy of Fallot	*NKX2–5*	600584
Atrial septal defect/ventricular septal defect	*TBX5*	601620
Pulmonary stenosis	*PTPN11*	176876
Pulmonary stenosis/tetralogy of Fallot	*JAG1*	601920
Tetralogy of Fallot/persistent truncus arteriosus, interrupted aortic arch	*?TBX1*	602054
Supravalvular aortic stenosis	*ELN*	130160
Isomerism/laterality defects	*ZIC3, CFC1*	300265, 605194
Patent ductus arteriosus	*TFAP2B, ZFHX1B*	601601; 605802

a further 5%. A substantial majority have no clear etiological basis (Kalter and Warnaky, 1983).

Non-disjunction of homologous chromosomes is the commonest mechanism leading to trisomies 13, 18 and 21 in humans. Post-zygotic loss of either an X or a Y chromosome leads to Turner syndrome which is characterized by a variety of cardiac, urological, skeletal and endocrine defects. More recently a host of disorders have been identified that are caused by recurrent stereotypic deletion of segments of DNA, mediated by long complex flanking repetitive sequences. Examples of these disorders include 22q11 deletion syndrome, Smith Magenis syndrome (chromosome 17p11.2) and Angelman/Prader-Willi (chromosome 15q11) syndromes. These contiguous gene deletion syndromes have been labeled ''genomic disorders'' in recognition of both the size and the recurrent deletional mechanism that leads to their appearance (Inoue and Lupski, 2002). It is possible that smaller stereotypic chromosomal deletions may underlie other as yet uncharacterized disorders, especially those that involve genes encoding proteins that are sensitive to alterations in gene dosage such as transcription factors. The intrinsically low reproductive fitness of many developmental disorders makes it difficult to undertake genetic studies to examine such hypotheses that invoke dominantly acting, reproductively lethal pathogenic mechanisms. Newer methodologies, such as comparative genomic hybridization, using microarrays may yield vital clues in this area.

Rapid progress has been made over the last 15 years in identifying the genes underlying many monogenic malformation syndromes giving clues to some of the genes critical for development of specific organs (Donnai and Read, 2003). As an example, genes known to underlie the pathogenesis of syndromic congenital heart disease are given in Table 13.3. Although numerically small in terms of their contribution towards morbidity and mortality, description of their pathogenesis has contributed disproportionately to the understanding of normal embryogenesis and has consequently set the stage for investigation of more complex disorders. The roles that epigenetic mechanisms play in the genesis of malformations remain a relatively unexplored and unappreciated area although they are implicated in disorders characterized by defects in imprinting (Paulsen and Ferguson-Smith, 2001) and underlying some developmental defects observed in animals cloned by somatic cell nuclear transfer techniques (reviewed in Li, 2002).

Robust epidemiological studies estimating the contribution that environment factors make towards congenital anomalies are lacking, but some data indicate that the contributions of

alcohol (May and Gossage, 2001) and maternal diabetes (Schwartz and Teramo, 2000) as teratogenic influences are considerable. Hampering attempts to quantify the effects of these agents is the observation that a given teratogen can have pleiotropic developmental effects and conversely many contributors, both environmental and genetic, can produce similar patterns of anomalous development. Adding to this complexity is emerging evidence to support the role of genetic factors in influencing the metabolism of, or degree of exposure to, human teratogens such as the anticonvulsant sodium valproate (Duncan *et al.*, 2001) and ethanol (Stoler *et al.*, 2002; Chambers and Jones, 2002).

Another newly recognized teratogenic influence is the transplacental passage of maternally derived antibodies directed against fetal neuromuscular proteins (Polizzi *et al.*, 2000). The resultant phenotypes are due to fetal akinesia and are broadly subsumed under the diagnosis of arthrogryposis multiplex congenita.

Together the emerging recognition of an interplay between certain susceptibility alleles and specific environmental factors has set the stage for a mechanistic explanation of the long established concept of liability and thresholds in the genesis of malformations in humans (Falconer, 1965). This thesis proposes that several genetic and environmental susceptibility elements must act in combinatorial fashion to raise an individual's liability to a threshold level beyond which a phenotype manifests. This theory appears consistent with the observed intrafamilial recurrence risks that relatives of a proband are predicted to exhibit for such ''polygenic'' traits (Edwards, 1960). Latterly others have contended that stochastic factors in concert with considerably fewer numbers of genes can exhibit the same characteristics (Kurnit *et al.*, 1987). Experimental data are largely lacking to support either model. The challenge that lies ahead is to define the nature of both environmental and genetic factors and their mode of interaction amidst much phenotypic and etiological heterogeneity.

Measuring the genetic contribution in developmental disorders

Strong evidence exists to point towards sizable genetic contributors to the genesis of congenital malformations in humans (Stevenson, 1993*a*). The two principal forms of clinical evidence which are traditionally used in this regard are the degree of familial aggregation that a trait exhibits, and twin concordance data. Differences in the prevalence of a disorder between sexes such as the observed male preponderance in pyloric stenosis (Carter and Evans, 1969) is also viewed as suggestive evidence for a genetic contribution. In the instance of congenital abnormalities, these measures have distinct limitations as tools to measure the heritability of these conditions.

Family studies carry the caveat that sibs share the same maternal environment over precisely the period during which the etiological determinants are exerting their effect i.e. during embryogenesis. Additionally, family members share environmental and socioeconomic status that contribute confounding effects in a transgenerational manner. This reduces the ability to distinguish between genetic and intrauterine environmental factors as causative influences. Twin concordance studies ascertain the difference in concordance of a trait between dizygotic and monozygotic twin pairs and this is interpreted as a measure of the genetic contribution towards aetiology. Such studies are limited by the observation that monozygotic twinning per se is associated with an increase in the prevalence of congenital anomalies (Schinzel *et al.*, 1979; Hall, 1996).

Epidemiological methods that can be used to deduce etiological factors underlying developmental disorders are therefore secondary and limited. Prevalence studies designed to detect changes over time, between geographical regions and over seasons of the year, act as possible indicators of environmental influences. Examples are studies indicating geographical gradients and seasonality (Fraser and Gwyn, 1998) as influences on the prevalence of facial clefting, and parallel alterations

Table 13.4. Sibling recurrence risk for short and long segment Hirschsprung disease

Type of involvement in proband	Male proband % affected		Female proband % affected	
	Brothers	Sisters	Brothers	Sisters
Short segment	4.3	0.9	2.0	–
Long segment	9.3	6.5	6.7	7.4
Overall	5.0	1.6	3.7	2.7

Modified from Garver *et al.* (1985).

in the incidence of pyloric stenosis and maternal smoking over time (Sorensen *et al.*, 2002).

Proxy measures of genetic influences, such as variation in the incidence of conditions between different ethnic populations, could be explained on the grounds of socioeconomic or confounding cultural influences and therefore their use must be interpreted with caution.

Evidence for a genetic contribution towards developmental disorders with complex etiology has also been obtained through laboratory studies examining animal and human subjects. Studies of monogenic malformation traits in mice have repeatedly shown that expressivity and penetrance varies with genetic background (Sibilia and Wagner, 1995; Threadgill *et al.*, 1995; Qu *et al.*, 1998). Dissection of the genetic background using congenic analysis and other approaches holds the promise that these lesser determinants of phenotypic expression may be identified and their individual contribution towards the heritability of these traits quantified (Southard-Smith *et al.*, 1999; Nadeau, 2003).

Examples of multiple loci interacting in the human to produce developmental disorders have been more difficult to identify. Certainly the variability recognized in Mendelian conditions could conceivably represent influences from environmental, stochastic and/or epistatic genetic effects. Variable expressivity of dominantly inherited developmental conditions, such as holoprosencephaly a disorder characterized by failure of midline cleavage of the forebrain (Roessler and Muenke, 2003), illustrate this point. This condition is also an example of disorders with high heritability and a limited number of contributing genes; the so-called oligogenic diseases (Ming and Muenke, 2002). Instances of their simplest forms – digenic diseases caused by the interaction of two loci – have been forthcoming in recent years (Katsanis *et al.*, 2001; Ming and Muenke, 2002; Slavotinek and Biesecker, 2003). The human developmental disorder with perhaps the most defined set of oligogenic contributors to date is a condition known as Hirschsprung disease where innervation of the lower intestinal tract is incomplete.

Identifying the multiple epistatic determinants underlying a developmental disorder: Hirschsprung disease

Hirschsprung disease (HSCR) is a condition characterized by hypomotility of the large intestine attributable to the absence of intramural enteric ganglia. The clinical consequences are constipation, intestinal obstruction and life-threatening enterocolitis in newborns. The condition can be divided into a short segment form where lack of innervation only affects the distal large intestine; and a long segment form characterized by absence of ganglia over more extensive regions of the colon. The length of affected intestine is correlated to the sib recurrence risk and inversely proportional to the ratio of affected males to females (Table 13.4; Garver *et al.*, 1985; Badner *et al.*, 1990).

This condition represents a paradigm for the dissection of a complex and relatively common

developmental disorder (present in 1 in 4000 liveborn infants) for several reasons. Firstly, studies on the heritability (defined as the degree to which genetic factors can explain the phenotypic variance of a condition) using twin data will probably be relatively free of confounding factors induced by the twinning process itself. Secondly, the phenotype is well described, exhibits familial clustering and has well-established criteria for diagnosis (Badner *et al.*, 1990). Increasing the attractiveness of this disorder as a model system are data indicating that a limited number of genes contribute to the genotype and that its heritability is at least 80%, an exceptionally high figure for a developmental disorder. These factors indicate that environmental influences are at best minor contributors to the well-established observations of variable expressivity and incomplete penetrance (Badner *et al.*, 1990).

A variety of approaches have been utilized to identify the genetic determinants underlying HSCR. Firstly, clinical and epidemiological studies established that long segment disease exhibited higher penetrance and less variable expressivity than short segment disease (Bodian and Carter, 1963; Garver *et al.*, 1985; Badner *et al.*, 1990). The implications for recurrence risk in family members reflect these observations (Table 13.4). The first gene of major effect in both short and long segment HSCR was identified as *RET* using classical linkage studies to map a deletion encompassing the *RET* locus in a familial case of the disease and then detection of mutations in isolated cases of HSCR. The characterization of other genes underlying syndromic HSCR has proved more challenging but several others have had mutations identified using similar approaches (Amiel and Lyonnet, 2001). A genome-wide association study in inbred populations with elevated susceptibility to the condition and work using animal models have identified epistatic relationships between some of these loci and still further uncharacterized genes (Carrasquillo *et al.*, 2002). Although studies replicating all of these interactions are awaited, these data indicate the feasibility of using variations of techniques previously utilized in the identification of genes underlying highly penetrant monogenic malformation syndromes. More recently, a genome-wide scan using an outbred population was adopted to study loci implicated in short segment disease which, although implicating the same gene of major effect, *RET*, identified different, unidentified contributory loci elsewhere in the genome (Bolk-Gabriel *et al.*, 2002).

The use of mouse models and genome-wide linkage screens to exploit the known high heritability of HSCR has enabled the definition of loci which may explain much, if not all of the risk for developing HSCR, in combinatorial fashion. Major challenges that remain are to identify the underlying genes and polymorphisms that confer these more minor genetic contributions to the HSCR phenotype and to provide some form of experimental verification that the proposed models of best fit (Bolk-Gabriel *et al.*, 2002) have a biological underpinning.

Beyond oligogenic disease

It remains to be seen if the application of similar approaches to define the genetic determinants of other developmental disorders will be as successful as for HSCR. Complicating issues may include identifying and controlling for environmental modifiers of the phenotype as well as developing methodologies to overcome combinatorial heterogeneity between contributing loci. Similar reliance upon studies utilizing inbred strains of mice may also be problematic with models that incorporate multiple interacting loci. There are many instances of disparities in penetrance between murine and human subjects in haploinsufficient states for even single developmental genes. In studying human subjects with congenital malformations, quantitative effects of gene function are difficult to measure (analogous to long versus short segment Hirschsprung disease) and will require innovative clinical approaches to classify phenotypic severity in order to construct plausible models on which to base gene discovery approaches. It may be

necessary to develop models using inbred strains of large domesticated animals (e.g. dogs, sheep) where it would be technically easier to study developmental processes in the early stages of embryogenesis while controlling for genetic background.

Identification of environmental contribution to developmental disorders

The difficulty involved in the identification of environmental contributors to developmental disorders is compounded by their diversity, and in turn by the multiplicity of malformations that can result. Much work has previously focused on establishing associations between particular environmental factors and congenital birth defects (reviewed in Stevenson, 1993*b*).

Classic teratogens

The ingestion of certain substances has been associated with fetal wastage and birth abnormalities since antiquity. A large number of teratogens have been identified by association through epidemiological means. Some of the effects of these agents are organ specific, the well-described association of the anti-nausea drug thalidomide with a specific form of underdevelopment of the limbs termed phocomelia, being a widely cited example (Smithells, 1973; Holmes, 2002).

A difficulty in assessing many of the reports of exposures to environmental agents associated with birth defects is the lack of a methodical strategy to quantitate the association and strength of the case for causality (Khoury *et al.*, 1991; Castilla *et al.*, 2001; Chapter 6, this volume). Improved surveillance and monitoring of the effects of new pharmaceuticals has led to the recognition of rarer, but nonetheless specific teratogenic side effects of more recently introduced agents. The full extent of the effects of the teratogenic influences of more established and readily available agents, such as alcohol, are still debated, but are undoubtedly significant. Congenital malformations have been associated with ingestion of excessive amounts of dietary factors such as vitamin A, but this uncommon scenario is unlikely to be a numerically significant determinant of the overall prevalence of malformations.

Physical influences, such as radiation and hyperthermia, have also been implicated in the genesis of developmental abnormalities. The overall magnitude of their contribution is likely to be small. Initial studies failed to detect an appreciable and consistent rise in the incidence of birth anomalies after the Chernobyl accident (Little, 1993) but this may have been due to problems with case ascertainment and accurate measures of exposure. It is intriguing to speculate that the acknowledged association between hyperthermia and birth anomalies (Graham and Edwards, 1998; Edwards *et al.*, 2003) may have a mechanistic foundation in the phenomenon of canalization which is now being understood in molecular terms (see below).

With the possible exception of alcohol, the contribution of all of these agents to the morbidity associated with developmental anomalies is likely to be small, purely on the basis of the size of the exposed population. In contrast, a theme that has emerged over the last decade is that nutritional deficiencies, combined with genetic predisposition, may be far weightier etiological influences.

Nutritional deficits as teratogenic influences

Evidence to support the thesis that deficient diet combined with defective intermediary metabolism is an underlying contributor to birth defects, first came from two large interventional trials which demonstrated that folic acid supplementation reduces the prevalence of neural tube defects (NTDs) by 70% (MRC Vitamin Study Research Group, 1991; Czeizel and Dudás, 1992). Anencephaly, encephalocoele, craniorachischisis, iniencephaly and spina bifida all fall within this spectrum of malformations. NTDs result when the neural folds fail to elevate and fuse in the dorsal

midline before the end of the fourth week post-conception. The prevalence of these conditions collectively is around 0.1% of all live births but varies markedly with ethnicity, socioeconomic status and geographical location (reviewed in Botto *et al.*, 1999). Twin concordance data and familial aggregation in humans indicate strong genetic contributors to their aetiology (reviewed in Frey and Hauser, 2003).

Subsequently work on the association between folate intake and NTDs has aimed to identify the biological processes which are modified by folate and also to institute public health interventions to increase consumption of folate by women during the periconceptual period by either voluntary means (Berry *et al.*, 1999) or via food supplementation (Ray *et al.*, 2002). These interventions have been shown to lead to an appreciable fall in the prevalence of NTDs (Ray *et al.*, 2002). These observations suggested the existence of genetic factors that predispose to structural malformation but which can be compensated for by pharmacological doses of folate. One such factor has been identified, a variant in the folate metabolizing gene *5,10 methylenetetrahydrofolate reductase* which accounts for 13% of the population attributable risk for neural tube defects (Whitehead *et al.*, 1995). Homozygosity of the fetus for the susceptibility allele confers an elevated risk of NTDs independent of maternal genotype, but possibly exacerbated by low maternal folate status (Shields *et al.*, 1999). The existence of mouse models of neural tube defects that can be partly or wholly compensated for by exogenous folate (Zhao *et al.*, 1996; Fleming and Copp, 1998; Carter *et al.*, 1999) and other substances (Greene and Copp, 1997), offer avenues by which other genes and dietary agents may be identified that contribute towards neural tube defects in humans (Jurlioff and Harris, 2000). Parallel studies have also suggested that folic acid supplementation may reduce the prevalence of other malformations such as orofacial clefts (Shaw *et al.*, 1995) and omphalocoele (Botto *et al.*, 2002).

The intrauterine diabetic environment: an intrauterine teratogenic milieu

The exposure of the developing embryo and fetus to metabolic derangements associated with maternal diabetes mellitus (DM) is a well documented, commonly occuring, but still poorly understood, example of a teratogenic environmental influence (Martinez-Frias, 1994; Reece *et al.*, 1996). A large number of structural malformations have been described in infants of diabetic mothers (Table 13.5). Despite an undisputed association of first trimester hyperglycaemia with the occurrence of malformations (Miller *et al.*, 1981), the teratogenicity of gestational diabetes, a transient state of glucose intolerance induced by pregnancy, is in dispute (Kalter, 1998; Kousseff, 1999). Recent studies indicate that the risk of malformations in infants of mothers with mild gestational DM is significantly lower than maternal diabetes established prior to conception (Sheffield *et al.*, 2002).

Table 13.5. Structural malformations associated with maternal diabetes mellitus

Central nervous system
Anencephaly/spina bifida
Holoprosencephaly
Cardiovascular system
Ventricular septal defects
Transposition of the great arteries
Abnormalities of cardiac situs
Craniofacial
Cleft lip/palate
Gastrointestinal
Small left colon syndrome
Anal/rectal atresia
Skeletal
Caudal regression syndrome
Sirenomelia
Femoral dysplasia
Rib/vertebral anomalies

Adapted from Cohen, 1997.

The exact mechanism behind the teratogenicity of the maternal diabetic state is poorly understood but tight glycemic control has been correlated with improved neurodevelopment (Sells *et al.*, 1994) and fewer structural malformations (Ray *et al.*, 2001) implying that maternal hyperglycemia or factors associated with it are responsible.

Some malformations are vastly over-represented in infants of diabetics compared to the general population (Table 13.5). These observations imply that certain developmental processes are profoundly influenced by metabolic derangements occurring in the maternal diabetic state (Reece *et al.*, 1996). An understanding of the molecular correlates to this clinical observation is lacking although the description of genes implicated in the genesis of some of these malformations (Belloni *et al.*, 2000, Roessler and Muenke, 2003) and modulation of their expression in animal models of diabetic embryopathy (Pani *et al.*, 2002) may soon present added insights.

Maternal obesity during pregnancy has also been recently recognized as a risk factor for congenital malformations, specifically NTDs, omphalocoele, heart defects and multiple anomaly states (Prentice and Goldberg, 1996; Watkins *et al.*, 2003). The rising prevalence of obesity possibly places this factor high on the list of numerically important teratogenic influences.

A recent retrospective study suggests that periconceptual vitamin supplementation may reduce the risk of malformations in the fetus of the diabetic mother (Correa *et al.*, 2003). This finding raises the possibility that dietary supplementation may prevent a proportion of the congenital malformations within this at-risk subgroup in a manner analogous to the observed reduction in NTDs by folate administration. Similarly these data may indicate the existence of genetic determinants modifiable by dietary factors which underlie susceptibility to malformation in the fetus exposed to the maternal diabetic intrauterine environment.

Approaches to identifying teratogens

Obvious ethical and logistical strictures will maintain epidemiological and animal studies at the forefront of efforts to identify harmful exogenous agents that contribute to human malformation. Evidence is amassing to suggest that the major environmental contributors include dysregulation of metabolism, perhaps exacerbated by dietary and lifestyle factors. The clear interactions between genetic and environmental factors in the causation of malformations in mice (Zhao *et al.*, 1996; Barbera *et al.*, 2002) may not translate directly to humans due to the outbred nature of human populations. It may be that, just as has been suggested for many supposed environmental carcinogens (Thilly, 2003), exogenous agents exert their teratogenic effect through the modulation of gene expression rather than by acting directly as mutagens.

Emerging evidence suggests that genetic regulatory networks possess an innate buffering mechanism, termed ''canalization'', which can be disturbed by both genetic and environmental influences, to unmask latent variability in phenotypes (Stearns, 2002). Examining this hypothesis will require a deeper understanding of the molecular basis of normal embryogenesis. The phenomenon of canalization potentially offers much in explaining how genetic and environmental factors confer susceptibility to developmental disorders.

Summary

Dissection of the genetic traits that underlie aetiologically complex diseases places reliance on assessing the heritability of the disease and modeling the number and strength of interacting loci. In the instance of developmental disorders this task is made doubly difficult by the limitations of the epidemiological tools at our disposal. Despite this much has been learnt about the genetic determinants underlying malformations. Progress in this area will continue to be made by utilizing lessons gained in the delineation of genes

of major effect in monogenic and oligogenic malformation syndromes and from studies of animal models where both the effects of environmental factors and genetic background can be controlled.

REFERENCES

Amiel, J. and Lyonnet, S. (2001). Hirschsprung disease, associated syndromes and genetics: a review. *J Med Genet*, **38**, 729–39.

Badner, J. A., Sieber, W. K., Garver, K. L. and Chakravarti, A. (1990). A genetic study of Hirschsprung disease. *Am J Hum Genet*, **46**, 568–80.

Barbera, J. P., Rodriguez, T. A., Greene, N. D. *et al.* (2002). Folic acid prevents exencephaly in Cited2 deficient mice. *Hum Mol Genet*, **11**, 283–93.

Belloni, E., Martucciello, G., Verderio, D. *et al.* (2000). Involvement of the *HLXB9* homeobox gene in Currarino syndrome. *Am J Hum Genet*, **66**, 312–19.

Berry, R. J., Li, Z., Erickson, J. D. *et al.* (1999). Prevention of neural-tube defects with folic acid in China. *N Engl J Med*, **341**, 1485–90.

Bodian, M. and Carter, C. O. (1963). A family study of Hirschsprung's disease. *Ann Hum Genet*, **26**, 261–77.

Bolk-Gabriel, S., Salomon, R., Pelet, A. *et al.* (2002). Segregation at three loci explains familial and population risk in Hirschsprung disease. *Nature Genet*, **31**, 89–93.

Bosi, G., Garani, G., Scorrano, M., Calzolari, E. and IMER Working Party. (2003). Temporal variability in birth prevalence of congenital heart defects as recorded by a general birth defects registry. *J Pediatr*, **142**, 690–8.

Botto, L. D., Moore, C. A., Khoury, M. J. and Erickson, J. D. (1999). Neural-tube defects. *N Engl J Med*, **341**, 1509–19.

Botto, L. D., Mulinare, J. and Erickson, J. D. (2002). Occurrence of omphalocele in relation to maternal multivitamin use: a population-based study. *Pediatrics*, **109**, 904–8.

Carter, C. O. and Evans, K. A. (1969). Inheritance of congenital pyloric stenosis. *J Med Genet*, **6**, 233–54.

Carter, M., Ulrich, S., Oofuji, Y., Williams, D. A. and Ross, M. E. (1999). Crooked tail (Cd) models human folate-responsive neural tube defects. *Hum Molec Genet*, **8**, 2199–204.

Carrasquillo, M. M., McCallion, A. S., Puffenberger, E. G. *et al.* (2002). Genome-wide association study and mouse model identify interaction between RET and EDNRB pathways in Hirschsprung disease. *Nat Genet*, **32**, 237–44.

Castilla, E. E., Lopez-Camelo, J. S., Campana, H. and Rittler, M. (2001). Epidemiological methods to assess the correlation between industrial contaminants and rates of congenital anomalies. *Mutat Res*, **489**, 123–45.

Chambers, C. D. and Jones, K. L. (2002). Is genotype important in predicting the fetal alcohol syndrome? *J Pediatr*, **141**, 751–2.

Cohen, M. M. (1997). *The child with multiple birth defects*. 2nd edn. Oxford, New York: Oxford University Press.

Correa, A., Botto, L., Liu, Y., Mulinare, J. and Erickson, J. D. (2003). Do multivitamin supplements attenuate the risk for diabetes-associated birth defects? *Pediatrics*, **111**, 1146–51.

Czeizel, A. E. and Dudás, I. (1992). Prevention of the first occurrence of neural tube defects by periconceptual vitamin supplementation. *N Engl J Med*, **327**, 1832–5.

Donnai, D. and Read, A. P. (2003). How clinicians add to knowledge of development. *Lancet*, **362**, 477–84.

Duncan, S., Mercho, S., Lopes-Cendes, I. *et al.* (2001). Repeated neural tube defects and valproate monotherapy suggest a pharmacogenetic abnormality. *Epilepsia*, **42**, 750–3.

Edwards, J. H. (1960). The simulation of Mendelism. *Acta Genet*, **10**, 63–70.

Edwards, M. J., Saunders, R. D., Shiota, K. (2003). Effects of heat on embryos and foetuses. *Int J Hyperthermia*, **19**, 295–324.

Falconer, D. S. (1965). The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Ann Hum Genet*, **29**, 51–71.

Fleming, A. and Copp, A. J. (1998). Embryonic folate metabolism and mouse neural tube defects. *Science*, **280**, 2107–8.

Fraser, F. C. and Gwyn, A. (1998). Seasonal variation in birth date of children with cleft lip. *Teratology*, **57**, 93–5.

Frey, L. and Hauser, W. A. (2003). Epidemiology of neural tube defects. *Epilepsia*, **44**, Suppl 3, 4–13.

Garver, K. L., Law, J. C. and Garver, B. (1985). Hirschsprung disease: a genetic study. *Clin Genet*, **28**, 503–8.

Graham, J. M. Jr. and Edwards, M. J. (1998). Teratogen update: gestational effects of maternal hyperthermia due to febrile illnesses and resultant patterns of defects in humans. *Teratology*, **58**, 209–21.

Greene, N.D. and Copp, A.J. (1997). Inositol prevents folate resistant neural tube defects in the mouse. *Nature Med*, **3**, 60–6.

Hall, J.G. (1996). Twinning: mechanisms and genetic implications. *Curr Opin Genet Dev*, **6**, 343–7.

Hoffman, J.I. and Kaplan, S. (2002). The incidence of congenital heart disease. *J Am Coll Cardiol*, **39**, 1890–900.

Holmes, L.B. (2002). Teratogen-induced limb defects. *Am J Med Genet*, **112**, 297–303.

Inoue, K. and Lupski, J.R. (2002). Molecular mechanisms for genomic disorders. *Annu Rev Genomics Hum Genet*, **3**, 199–242.

Jurlioff, D.M. and Harris, M.J. (2000). Mouse models for neural tube defects. *Hum Molec Genet*, **9**, 993–1000.

Kalter, H. and Warnaky, J. (1983). Medical progress. Congenital malformations: etiologic factors and their role in prevention. *N Engl J Med*, **308**, 424–31.

Kalter, H. The non-teratogenicity of gestational diabetes (1998). *Paediatr Perinat Epidemiol*, **12**, 456–8.

Katsanis, N., Ansley, S.J., Badano, J.L. *et al.* (2001). Triallelic inheritance in Bardet-Biedl syndrome, a Mendelian recessive disorder. *Science*, **293**, 2256–9.

Khoury, M.J., Beaty, T.H. and Cohen, B.H. (1991). Applications of the concept of attributable fraction in medical genetics. *Am J Med Genet*, **40**, 177–82.

Kouseff, B. (1999). Gestational diabetes mellitus (class A): a human teratogen? *Am J Med Genet*, **83**, 402–8.

Kurnit, D.M., Layton, W.M. and Matthysse, S. (1987). Genetics, chance and morphogenesis. *Am J Hum Genet*, **41**, 979–95.

Li, E. (2002). Chromatin modification and epigenetic reprogramming in mammalian development. *Nat Rev Genet*, **3**, 662–73.

Little, J. (1993). The Chernobyl accident, congenital anomalies and other reproductive outcomes *Paediatr Perinat Epidemiol*, **7**, 121–51.

Martinez-Frias, M.L. (1994). Epidemiological analysis of outcomes of pregnancy in diabetic mothers: identification of the most characteristic and most frequent congenital anomalies. *Am J Med Genet*, **51**, 108–13.

May, P.A. and Gossage, J.P. (2001). Estimating the prevalence of fetal alcohol syndrome. A summary. *Alcohol Res Health*, **25**, 159–67.

Miller, E., Hare, J.W., Cloherty, J.P. *et al.* (1981). Elevated maternal hemoglobin A1c in early pregnancy and major congenital anomalies in infants of diabetic mothers. *N Engl J Med*, **304**, 1331–4.

Ming, J.E. and Muenke, M. (2002). Multiple hits during early embryonic development: digenic diseases and holoprosencephaly. *Am J Hum Genet*, **71**, 1017–32.

MRC Vitamin Study Research Group. (1991). Prevention of neural tube defects: results of the Medical Research Council vitamin study. *Lancet*, **338**, 131–7.

Nadeau, J.H. (2003). Modifier genes and protective alleles in humans and mice. *Curr Opin Genet Dev*, **13**, 290–5.

Pani, L., Horal, M. and Loeken, M.R. (2002). Polymorphic susceptibility to the molecular causes of neural tube defects during diabetic embryopathy. *Diabetes*, **51**, 2871–4.

Paulsen, M. and Ferguson-Smith, A.C. (2001). DNA methylation in genomic imprinting, development and disease. *J Pathol*, **195**, 97–110.

Polizzi, A., Huson, S.M. and Vincent, A. (2000). Teratogen update: maternal myasthenia gravis as a cause of congenital arthrogryposis. *Teratology*, **62**, 332–41.

Prentice, A. and Goldberg, G. (1996). Maternal obesity increases congenital malformations. *Nutr Rev*, **54**, 146–50.

Qu, S., Tucker, S.C., Ehrlich, J.S. *et al.* (1998). Mutations in mouse Aristaless-like4 cause Strong's luxoid polydactyly. *Development*, **125**, 2711–21.

Ray, J.G., Meier, C., Vermeulen, M.J. *et al.* (2002). Association of neural tube defects and folic acid food fortification in Canada. *Lancet*, **360**, 2047–8.

Ray, J.G., O'Brien, T.E. and Chan, W.S. (2001). Preconception care and the risk of congenital anomalies in the offspring of women with diabetes mellitus: a meta-analysis. *QJM*, 94, 435–44.

Reece, E.A., Homko, C.J. and Wu, Y.K. (1996). Multifactorial basis of the syndrome of diabetic embryopathy. *Teratology*, **54**, 171–82.

Roessler, E. and Muenke, M. (2003). How a hedgehog might see holoprosencephaly. *Hum Mol Genet.* **12**, R15–R25.

Schinzel, A.A., Smith, D.W. and Miller, J.R. (1979). Monozygotic twinning and structural defects. *J Pediatr*, **95**, 921–30.

Schwartz, R. and Teramo, K.A. (2000). Effects of diabetic pregnancy on the fetus and newborn. *Semin Perinatol*, **24**, 120–35.

Sells, C.J., Robinson, N.M., Brown, Z. and Knopp, R.H. (1994). Long-term developmental follow-up of infants of diabetic mothers. *J Pediatr*, **125**, S9–S17.

Shaw, G.M., Lammer, E.J., Wasserman, C.R., O'Malley, C.D. and Tolarova, M.M. (1995). Risks of orofacial clefts in children born to women using multivitamins

containing folic acid periconceptionally. *Lancet*, **346**, 393–6.

Sheffield, J. S., Butler-Koster, E. L., Casey, B. M., McIntire, D. D. and Leveno, K. J. (2002). Maternal diabetes mellitus and infant malformations. *Obstet Gynecol*, **100**, 925–30.

Shields, D. C., Kirke, P. N., Mills, J. L. *et al.* (1999). The ''thermolabile'' variant of methylenetetrahydrofolate reductase and neural tube defects: an evaluation of genetic risk and the relative importance of the genotypes of the embryo and the mother. *Am J Hum Genet*, **64**, 1045–55.

Sibilia, M. and Wagner, E. F. (1995). Strain-dependent epithelial defects in mice lacking the EGF receptor. *Science*, **269**, 234–8.

Slavotinek, A. and Biesecker, L. G. (2003). Genetic modifiers in human development and malformation syndromes, including chaperone proteins. *Hum Mol Genet*, **12**, R45–R50.

Smithells, R. W. (1973). Defects and disabilities of thalidomide children. *Br Med J*, **1**(5848), 269–72.

Sorensen, H. T., Norgard, B., Pedersen, L., Larsen, H. and Johnsen, S. P. (2002). Maternal smoking and risk of hypertrophic infantile pyloric stenosis: 10 year population based cohort study. *BMJ*, **325**, 1011–12.

Southard-Smith, E. M., Angrist, M., Ellison, J. S. *et al.* (1999). The Sox10 (Dom) mouse: modeling the genetic variation of Waardenburg-Shah (WS4) syndrome. *Genome Res*, **3**, 215–25.

Spranger, J., Benirschke, K. and Hall, J. G. *et al.* (1982). Errors of morphogenesis: concepts and terms. *J Pediatr*, **100**, 160–5.

Stearns, S. C. (2002). Progress on canalization. *Proc Natl Acad Sci USA*, **99**, 10229–30.

Stevenson, R. E. (1993*a*). The genetic basis of human anomalies. In Stevenson, R. E., Hall, J. G. and Goodman, R. M. eds., *Human malformations and related anomalies, vol 1*. New York: Oxford University Press, pp. 115–35.

Stevenson, R. E. (1993*b*). The Environmental Basis of Human Anomalies. In Stevenson, R. E., Hall, J. G. and Goodman, R. M. eds., *Human malformations and related anomalies, vol 1*. New York: Oxford University Press, pp. 137–68.

Stoler, J. M., Ryan, L. M. and Holmes, L. B. (2002). Alcohol dehydrogenase 2 genotypes, maternal alcohol use, and infant outcome. *J Pediatr*, **141**, 780–5.

Thilly, W. G. (2003). Have environmental mutagens caused oncomutations in people? *Nat Genet*, **34**, 255–9.

Threadgill, D. W., Dlugosz, A. A., Hansen, L. A. *et al.* (1995). Targeted disruption of mouse EGF receptor: effect of genetic background on mutant phenotype. *Science*, **269**, 230–4.

van Essien, F. B. and Wannberg, S. L. (1993). Methionine but not folinic acid or vitamin B-12 alters the frequency of neural tube defects in *Axd* mutant mice. *J Nutrition*, **123**, 27–34.

Watkins, M. L., Rasmussen, S. A., Honein, M. A., Botto, L. D. and Moore, C. A. (2003). Maternal obesity and risk for birth defects. *Pediatrics*, **111**, 1152–8.

Whitehead, A. S., Gallagher, P., Mills, J. L. *et al.* (1995). A genetic defect in 5,10 methylene tetrahydrofolate reductase in neural tube defects. *Q J Med*, **88**, 763–6.

Wilcox, A. J., Baird, D. D. and Weinberg, C. R. (1999). Time of implantation of the conceptus and loss of pregnancy. *N Engl J Med*, **340**, 1796–9.

Yang, Q., Khoury, M. J. and Mannino, D. (1997). Trends and patterns of mortality associated with birth defects and genetic diseases in the United States, 1979–1992: an analysis of multiple-cause mortality data. *Genet Epidemiol*, **14**, 493–505.

Zhao, Q., Behringer, R. R. and de Crombrugghe, B. (1996). Prenatal folic acid treatment suppresses acrania and meroanencephaly in mice mutant for the *Cart1* homeobox gene. *Nature Genet*, **13**, 275–83.

14

Genes, environment and cancer

D. Timothy Bishop

Introduction

Epidemiological studies have shown that environmental exposures influence a person's risk of disease. For instance, migration studies show that a person's risk of cancer changes as he/she moves across environments; such changes must be due to changing exposures, as the host genome remains essentially constant (Haenszel, 1982). Studies comparing the lifestyle of persons with cancer (cases) and those without cancer (controls) have identified a number of consistent differences, assumed to be risk factors for that cancer, although the critical aspect of exposure is usually unknown or poorly measured. For instance, case-control studies show that persons with higher levels of smoking are more at risk of lung cancer, or those reporting diets higher in animal fat have a higher risk of bowel cancer but the critical combination and timing of mutagens, carcinogens, tumor promoters etc. is unknown. Instead, exposures are reported as number of packs of cigarettes per year or estimated average number of calories from fat. Chapter 11 discusses some of these issues in greater depth including the definition of the term ''environment''.

Genetic studies have shown that, for some persons, susceptibility is determined at least in part by genes (Eeles *et al.*, 2004). Family history studies indicate that for most common cancers, close relatives of cases have a risk of that same cancer which is between twice and four times that of the general population (Goldgar *et al.*, 1994; Hemminki *et al.*, 2004). Genes associated with high risks of cancer (penetrance) have been identified for a number of common cancers (Eeles *et al.*, 2004). For the majority of such susceptibilities and the majority of populations, the risk of cancer is high but the carrier frequency is low. For breast cancer in the UK, the carrier frequency of *BRCA1* mutations is estimated to be 1 in 1000 people (although higher for persons of Ashkenazi descent) (Peto *et al.*, 1999). As such, among non-Ashkenazim, *BRCA1* and *BRCA2* are estimated to explain less than 20% of the family aggregation of that cancer among young women, and less at older ages (Peto *et al.*, 1999). The remaining family aggregation may be due to other genes (Antoniou *et al.*, 2002), to the sharing of genes and environment or to the sharing of environmental exposures (risk factors) among relatives. Analysis of family aggregation of environmental exposures suggest that few exposures are strong enough or sufficiently shared among relatives to explain the observed family aggregation although, of course, the relative contribution of each explanation will differ by anatomical site and exposures (Hopper and Carlin, 1992).

In cancer epidemiology, the study of genes and environment seeks to rationalize the observations of the role of the two forms of exposure and to understand their interplay. Clearly, there will be a major interest in situations in which the presence of the genetic susceptibility has a qualitative effect on the nature of the risk associated with the environmental exposure, such as when persons

with particular genotypes have a greatly different pattern of risk by environmental exposure to that of other genotypes. In medicine, some examples are extreme. Persons with xeroderma pigmentosum (XP) have an extreme sun sensitivity in early childhood and a strong tendency to develop multiple malignant skin tumors after exposure to ultraviolet C irradiation (Kraemer and Slor, 1985; Arlett and Lehmann, 2004). This is now known to be due to defects in excision repair for the majority of cases so, while there is an association between sun exposure and the occurrence of skin tumors in the general population, the effect is far stronger in those persons with XP.

Studies to investigate genes and environment

The exposures that could be considered in this context are varied and could be an exposure such as radiation, cigarette smoke or a virus, or a behavior such as late age at first pregnancy. Consider a simplistic table (Table 14.1) in which persons may be genetically susceptible (with genotype) or have a more limited susceptibility (without genotype), and exposed to an environmental factor or not: most situations of interest are more complicated, involving multiple genes and/or genotypes plus exposures which may vary in strength and timing between persons. In Table 14.1, there are then 4 types of persons; each person fits into exactly one group depending on their genes and exposure. Persons in these four groups may have inherently different risks of disease; we denote these by R1, R2, R3 and R4. The magnitude of these four risks characterizes the joint effect of genes and environment. Typically, in epidemiology, one of the four risks is taken as a base-line (usually, the ''without genotype/Not exposed'' group whose risk of disease R1). The risks of disease in the other three groups are then compared to the base-line risk. So, R4/R1 is the relative risk of disease in those with both the genotype and the exposure as compared to the risk in people without either genotype or environmental exposure.

Table 14.1. Persons are classified into one of four categories depending upon their genotype and environmental exposure. For each of these four categories, there is an inherent risk of disease. More realistic examples are likely to involve risks varying by extent of environmental exposure and by precise genotype

	Not Exposed	Exposed
Without genotype	R1	R2
With genotype	R3	R4

Biologically we are interested in the magnitude and the pattern of risks (R1–R4). Biologically plausible patterns include (but are not limited to) (i) the situation in which both a genetic susceptibility and an environmental exposure are required for disease (e.g. as for phenylalanine in the diet and phenylalanine hydroxylase deficiency for phenylketonuria as the disease), (ii) the pattern of risks described above for XP, (iii) the situation in which both genes and environment are associated with an increase in risk and the combined effect of the two is greater than the effect of each separately. If we have in mind a causal mechanism for the joint effect of gene and environment then the interpretation of the relative strength of R1–R4 would be made in the light of this model.

Statistical tests can be developed to test specific hypotheses about the relationship between the separate effect of genes and environment on risk but they make extremely specific assumptions about the form of the interplay. The most natural hypothesis to test is to detect the independence of the two factors under a multiplicative model. Under that model, the combined risk of genes and environment is the product of the risk attributable to genes and the risk attributable to environment. Thus if having the genetic susceptibility trebles risk and having the environmental susceptibility doubles risk the relationship between R1:R2:R3:R4 should be in the proportions

1:2:3:6 ($=2\times3$): that is, there should be a sixfold difference in risk between having both the genetic susceptibility and the environmental exposure, as compared to having neither. The statistical test focuses on a specific hypothesis of independence but depends upon the precise scale of measurement employed to define the hypothesis; lack of independence is termed interaction. Under differing scales of measurement, conclusions of independence and interaction may then differ. As biological hypotheses are not usually so precise on a statistical level, the statistical test of independence (or interaction) is often of only limited interest. Further, showing that the risks followed a pattern of 1:2:3:6 would indicate lack of a statistical interaction but biologically would show the interplay of genes and environment.

The Melanoma Genetics Consortium (Bishop *et al.*, 2002) examined the risk (penetrance) for melanoma among *CDKN2A* mutation carriers. Carriers of particular germline mutations in *CDKN2A* are known to be at increased risk of melanoma. In this analysis, based on families with multiple persons with melanoma and a germline *CDKN2A* mutation, the overall risk of melanoma was estimated to be 62% by age 75 years. However, there was statistical evidence of a difference with carriers in Australia having a risk higher than that of the USA which, in turn, was higher than that in Europe. In the general population, for comparable time periods, the risk of melanoma by age 75 was at its highest in Australia in which it approached 2%. On an age-specific basis, the population rates in Australia were approximately 7 times higher than in Europe, while the rates in USA were 5 times higher than those in Europe. With the limited precision of the estimates of the penetrance of *CDKN2A* in the published analysis, the data are consistent with the age-specific *CDKN2A* penetrance rate being in the same proportions across these populations as melanoma in the general population (the majority of whom do not carry a *CDKN2A* mutation). Thus there is no evidence of a statistical interaction under a multiplicative model but biologically the data are consistent with the same factors affecting *CDKN2A* penetrance for melanoma as affect the risk of melanoma in the general population.

Studies to compare or estimate R1–R4 are required to determine the joint effects of genes and environment. If we knew precisely the relevant gene and precisely how to measure the environmental exposure (and this also could be determined early in life) then persons could be classified into one of the four categories and the joint effects estimated by following these four groups of people to estimate and compare the risks of the outcome of interest. Of course, studies of interest (termed ''cohort studies'', especially for a cancer endpoint, would be long lasting and require the identification of substantial groups of people at the outset. More realistically, we examine the relative values of R1–R4 by comparing the distribution of cases and controls for the 2×2 table defined above. Cases and controls will again fit into one of the four categories described above and the relative frequencies of the 4 groups between cases and controls gives information on the relative values of R1–R4. For such studies in which extrapolation of the results to the general population is necessitated, care must be taken to ensure that cases are representative of the total population of cases and that controls are representative of the general population which contains the cases. Any selection bias of cases or controls limits the generalizability of the findings; any studies which focus in the design on removing such potential biases is termed ''population-based''.

Study design

For case-control studies, practical and efficient study designs for the investigation of the joint effects of genes and environment are required. Case selection is usually straightforward, with the availability of Cancer Registries or, other population-based resources, listing diagnoses by geographical region and date. Access to such lists varies, depending on Data Protection rules etc. However, often, for case recruitment into

epidemiological studies, the only notable concern is the timing of the approach to the case, which requires a balance between imposing on subjects at the stressful time of diagnosis, with the concern that case mortality or extreme morbidity may impact on the representativeness of the cases included in the study.

Control recruitment is, however, considerably more challenging. Population-based control collection requires identifying and contacting people identified from appropriate lists or registers (such as primary care physician (GP) lists in the UK, electoral rolls in Australia) with particular characteristics (predominantly age, gender, residence, race or ethnicity). Access to such lists is likely to be more restrictive than access to case registers (Data Protection etc.) but contact is even more challenging, given the extent of telemarketing and unrequested mail, which competes for attention with a request to enter a study. Lack of representativeness is therefore more of an issue and less quantifiable than for cases. Finally, one technical issue involves the potential for the controls to come from a different gene pool to the case individuals. This could create spurious genetic associations, as variants which differed between the population which contains the case individuals and that for the controls would be interpreted as being causal for disease, rather than being a measure of the characteristics of the populations (Ewens and Spielman, 1995). The evidence for such unrecognized population stratification is limited but the issue remains a concern. For this variety of reasons, attention has begun to focus on the choice of other family members as controls in such studies (Gauderman, Witte and Thomas, 1999; Goldstein and Andrieu, 1999). The ideal choice would be siblings of a similar age to the case individual. Such studies have the potential for improving recruitment by targeting a population who may be more likely to participate, although, for ethical studies, care must be made not to put pressure on case individuals to contact relatives or on relatives to enrol as controls and to participate simply because of their relatedness to the case individuals. Also, statistically they are sometimes less efficient, as controls are related to case individuals implying that each relative provides less information about the general population; the price for this reduction in efficiency is increased study size (Gauderman, Witte and Thomas, 1999, Goldstein and Andrieu, 1999). Such designs have been termed case-control-family designs (Hopper, 2003).

Environment and high penetrance genes

Genes associated with high penetrance of cancer have been identified for breast cancer susceptibility (genes *BRCA1, BRCA2* and *TP53*) (Chapters 15,16), colorectal cancer (APC germline mutations and mismatch repair genes, predominantly *hMSH2* and *hMLH1*) (Chapter 17) and melanoma (*CDKN2A*) (Eeles *et al.*, 2004). The risks of malignancy at one or more anatomical site are 0.60 or higher by age 75 years for mutations in each of these genes. Population-based studies suggest that approximately 1 in 500 of the general population carries a germline mutation in *BRCA1* or *BRCA2* (although some populations, such as Ashkenazim, have notably higher carrier frequencies. Carrier frequencies for germline mutations in the other predisposing genes may be marginally lower than for *BRCA1* and *BRCA2*. Persons carrying such mutations represent a challenge for clinical management and an important group in which to examine the effects of environmental exposures as any exposures which are modifiable would be targets for manipulation. A number of approaches have been taken for examining the association of exposures with disease among carriers.

Typically, the study design involves examining the association of exposure with malignancy status among persons carrying a mutation. Practically, the limited numbers of mutation carriers known to date imposes major design and size constraints on the epidemiological studies. As there is no population screening for mutation carriers, the persons involved in these studies come from families with multiple cases of malignancy whose germline predisposition has been recognized. Such families

cover multiple generations, birth cohorts, clinical management schemes and healthcare systems so that comparisons between affected and unaffected carriers are not straightforward and interpretation is difficult. For instance, knowledge of carrier status among family members will depend upon which family members choose to have genetic testing; such choices may well depend upon factors such as parity and the interest in acquiring this knowledge for family rather than entirely personal reasons. These confusions about the actual structure of the data may explain why there have been such inconsistencies between studies. For instance, among *BRCA1* and *BRCA2* mutation carriers, studies of smoking have suggested that smoking influences risk (Brunet *et al.*, 1998) but a subsequent, larger study by the same investigators found no such effect (Ghadirian *et al.*, 2004). For these studies, attempts to make the studies population-based have been discussed as a means to removing such inadequacies in data collection (Hopper, 2003).

Over time, as numbers increase, better studies can be conducted to look at the effects of exposure or to examine interventions. For instance, CAPP2 is a chemoprevention study of neoplasia among persons with a germline mutation in a mismatch repair gene associated with an increased risk of bowel neoplasia. Participants are randomized to a treatment of aspirin, or an inactive placebo plus resistant starch, or an inactive placebo, as aspirin has been shown to reduce the risk of bowel neoplasia in various settings while epidemiological evidence suggests that resistant starch may be an important dietary difference between populations with differing rates of colorectal cancer.

Investigations of specific cancers

Bowel cancer

The descriptive epidemiology of bowel cancer indicates wide variation in cancer incidence rates across the world with the highest rates in Western Europe, USA, Canada, Australia and New Zealand (Potter and Hunter, 2002). The lowest rates are in parts of South Africa and Southern Asia. Rates have varied marginally by time across most countries but the most notable change has been in Japan where the incidence rate have risen from being among the lowest to among the highest in a couple of generations presumed to be due to the Westernization of diet and lifestyle in that time period. Diet is regarded as being the most important factor influencing cancer risk. A number of aetiological studies have examined individual differences in diet and lifestyle, especially between persons with and without cancer, either through cohort or case-control studies. Among dietary factors, diets high in animal fat and low in fiber, fruit and vegetables are associated with increased risks (Potter and Hunter, 2002). The complexity of diet assessment and the interrelationship of dietary items exacerbates the difficulty of identifying the critical aspects of diet. Several hypotheses have therefore been proposed.

Meat, heterocyclic amines and risk

Epidemiological studies have indicated that meat intake is a risk factor for colorectal cancer. These studies have been varied in design (cohort and case-control), size, population studied, methodology for assessing meat consumption and in their conclusions (reviewed by Potter *et al.*, 1993). Some studies have suggested that red meat is particularly associated with risk of this cancer while others find no effect. Some studies have found protective effects of poultry and fish consumption while the findings of others have not confirmed this. A recent European study, which represents probably the most thorough investigation of this issue to date, found both an increased risk from red and processed meat and a protective effect of fish (Norat *et al.*, 2005). Similar comments of inconsistency can be made about the association between cigarette smoking and colon cancer risk. As qualitatively, if not quantitatively, exposure to tobacco is more readily measured than quantity

and preparation style of meat, the inconsistency of the results is surprising. One speculation on the basis of these studies is that members of the public are differentially susceptible to the effects of such exposures.

One of the primary hypotheses concerns interindividual variation in the handling of heterocyclic amines (HAs). HAs are formed when animal protein is cooked at high temperature and are also present in cigarette smoke. Also, HAs are known to be carcinogens in rodents. There are at least three polymorphic enzymes involved in this process, CYP1A2, NAT1 and NAT2 (Kadlubar *et al.*, 1992). To date, genetic polymorphisms have been found in NAT1 and NAT2 which are associated with phenotypic differences (defined as ''fast'' or ''slow'' depending on the activity of the alleles) in their metabolism. Indeed, several studies have suggested that the fast phenotype of N-acetyltransferase NAT2 may confer susceptibility to colorectal cancer especially in individuals with a diet high in heterocyclic amines such as is contained in well-done meat (see below). Such a relationship could be postulated in terms of NAT2's role in activating such dietary heterocyclic amines.

There are two hypotheses as to how NATs could relate to colorectal cancer risk. The first mechanism involves CYP-mediated N-hydroxylation of arylamines which yields electrophilic intermediates that can be inactivated by NAT acetylation. The other mechanism suggests that fast acetylators are at increased risk because of the activation of procarcinogens such as HAs. Exposure to HAs is common; such HAs are produced when meat and fish are cooked at temperatures achievable in the household or by smoking. The HAs undergo hepatic N-oxidation and N-glucuronidation resulting in conjugated N-hydroxy metabolites which can be transported to the colonic lumen. Within the mucosa, these derivatives can be O-acetylated and can form covalent DNA adducts.

In a small Australian case-control study, Roberts-Thomson *et al.* compared colorectal cancer cases and controls (Roberts-Thomson *et al.*, 1996). The fast acetylator phenotype of NAT2 was associated with 1.8-fold increased risk of colorectal cancer with the highest risk occurred in the youngest group (<64 years of age) of cases. The risk of cancer increased with increasing intake of meat in fast but not in slow acetylators (see Chapter 11).

Other subsequent, marginally larger epidemiological studies have not confirmed precisely the findings of the Australian study (Roberts-Thomson *et al.*, 1996). A case-control study in the North of England found no difference in the frequency of the fast acetylator genotype between cases and controls (Welfare *et al.*, 1997). There was, considerable heterogeneity in dietary risk factors between fast and slow acetylators. The combination of fast acetylator status and frequent fried meat consumption in cases had an estimated sixfold increased risk of cancer. In the Physicians' Health Study study in the USA, again there was no overall independent association of NAT acetylation genotypes and colorectal cancer risk (Chen *et al.*, 1998). A stronger association of red meat intake with cancer risk among NAT rapid acetylators, especially among men 60 years or older was, however, identified. Among those men who were rapid acetylators for both *NAT1* and *NAT2*, consumption of >1 serving of red meat per day was associated with an estimated sixfold increased risk as compared with consumption of at most 0.5 servings per day. These prospective data are consistent with the suggestion that polymorphisms in the *NAT* genes confer differential susceptibility to the effect of red meat consumption on colorectal cancer risk although the analyses differ in the precise risk factors considered and found to be discriminating.

Among the larger studies, a further study in the North of England and Scotland found no evidence of variation in risk attributable either to the *NAT1* or *NAT2* polymorphisms or to meat intake (Barrett *et al.*, 2003). In a case-control study conducted in the USA, the amount of red and white meat consumed was not associated with overall colon cancer risk (Kampman *et al.*, 1999). Processed meat

consumption was weakly positively associated with colon cancer risk in men only (odds ratio for highest versus lowest quintile of intake = 1.4, 95% confidence interval = 1.0–1.9). No significant associations with colon cancer risk were observed for different *NAT2* gene variants although there were trends in the data for the fast acetylator group.

Slattery *et al.* (1998) found that *NAT2* polymorphisms were not significantly associated with colon cancer, except among older women, in whom the intermediate/fast phenotype was associated with a marginally increased risk (about 40%) of colon cancer. Cigarette exposure was associated with colorectal cancer risk with those smoking a pack or more of cigarettes per day being at an approximately 40% increased risk of colon cancer. No significant interaction between the genotypes and cigarette smoking and colon cancer was found.

This example shows the complexity of testing a biological hypothesis. The limited understanding of the details of the biological pathways, the lack of knowledge of the clearest measurement of exposure, and the difficulties of measuring that exposure all limit the interpretation. The positive findings are of some interest, and of some limited consistency, but clearly the various studies cannot be regarded as complete replications.

Folate and colorectal cancer

A second hypothesis relates folate to bowel cancer risk. This hypothesis unites the observation of low intake and plasma levels of folate and increased use of alcohol from the epidemiological studies as risk factors for bowel cancer. It proposes that abnormalities in folate levels disrupt nucleotide synthesis (and hence biological methylation), as well as double strand breaks and reduces DNA stability (both features of colorectal cancer) (Ames, 1999; Vogelstein *et al.*, 1989). Folic acid is the universal methyl donor in cellular methabolism. Folate deficiency may therefore result in global hypomethylation, which is thought to be an early event in the development of many colorectal tumors.

Germline mutations in mismatch repair genes are associated with greatly increased risks of colorectal cancer and cancer at other sites in hereditary non-polyposis colorectal cancer syndrome (HNPCC). Mutation carriers have a notably early onset of colorectal cancer and their tumors display a characteristic phenotype, termed ''microsatellite instability'' (MSI). MSI-positive tumors show variant alleles at many microsatellite and mononucleotide repeat genetic sequences. Examination of case series have shown that up to 15% of tumors exhibit this phenotype. MSI-positive tumors are found at young ages, and are often associated with a strong family history. Such tumors have lost mismatch repair capabilities through the inactivation of both alleles, often by loss of heterozygosity of the wild-type allele. However, MSI-positive tumors are found at much greater frequency at older ages, particularly in right-sided tumors and are more common in females. Immunohistochemical analysis of colorectal cancer tumours shows loss of mismatch repair protein expression (either *hMLSH2* or *hMLH1*) at young ages but predominantly loss of *hMLH1* at older ages. This loss is now known to be primarily due to promotor hypermethylation of the *hMLH1* gene, leading to defective mismatch repair capabilities (Cunningham *et al.*, 1998). Figure 14.1 shows how loss of expression of *hMSH2* and *hMLH1* relates to age and gender.

The possibility exists that abnormalities in methylation contribute to colorectal cancer development and that this could involve either folate in the diet, genetic polymorphisms or both. The polymorphism that has been studied in most detail involves the C677T variant of the methylenetetrahydrofolate reductase (*MTHFR*), which is a regulatory enzyme in the metabolism of folate. The variant T-allele produces a protein with decreased activity. Several case-control studies have found that the TT genotype has a reduced risk of colorectal cancer although these comparisons with the wild-type (CC) genotype have been of marginal significance. The reduced risk is most notable at higher levels of folate intake,

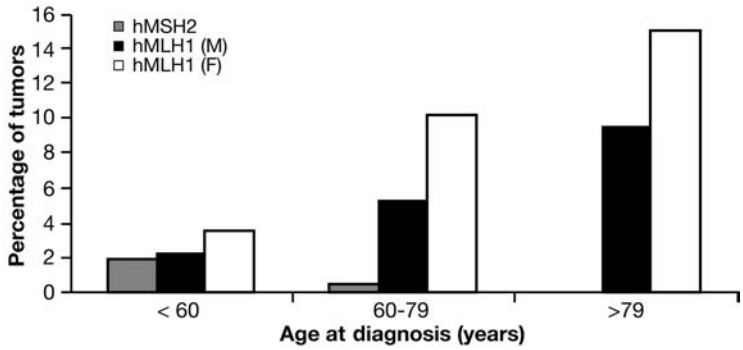

Figure 14.1 The percentage of colorectal tumors showing loss of expression of two mismatch repair proteins, hMSH2 and hMLH1 in a population case series of 730 cases (Coggins, unpublished data). The results are presented by gender for hMLH1 showing the differences between males (M) and females (F). Small numbers preclude such an analysis for hMSH2.

with suggestions that the effect is only found in those drinking low amounts of alcohol.

Finally, in a case-control study of smoking and colorectal cancer, the effect of smoking was estimated to be responsible for 21% of the tumors with microsatellite instability suggesting that smoking may have an effect on a pathway influencing or involving methylation (Slattery *et al.*, 2000).

Melanoma

Rates of melanoma vary widely geographically, with the major determinant of risk being skin color. The relative incidence of subtypes of melanoma differs across racial groups with one rare subtype termed ''acral lentiginous melanoma'' occurring in all racial groups (Hudson and Krige, 1993; Stevens *et al.*, 1990; Jimbow and Kukita, 1984) but the commoner subtypes (superficial spreading melanoma, nodular melanoma and lentigo maligna) occur essentially only in white persons. For these latter subtypes and among white populations, the highest incidence is in countries nearest to the equator, leading to the hypothesis that sun exposure is the major environmental determinant of melanoma risk (Armstrong, 1988). Australia and New Zealand have the highest rates worldwide (Jones *et al.*, 1999; Parkin *et al.*, 1997).

Variation in skin color explains much of the variation within and between geographical locations. For instance, in Hawaii, the melanoma incidence rates for people of European descent are much higher than for people of Polynesian descent (Chuang *et al.*, 1999) while within Europe, rates are higher in Scandinavian countries and the Alps than in southern Mediterranean countries (Parkin *et al.*, 1997).

Age-specific melanoma incidence rates across the world show two different patterns of risk. For high incidence regions, the pattern of the rates is similar to that for most common cancers; that is, the incidence is low at young ages but then rates increase exponentially (the familiar log–log relationship). In populations with lower rates, the rates are more evenly spread across age groups. This has led to the hypothesis that there are two different diseases, one associated with acute, the other with chronic exposure (Rivers, 2004).

MC1R, *coloration and melanoma risk*

Genes which determine skin colour are clear candidates for modifying the effect of UV exposure with respect to melanoma risk. While these genetic determinants are not known precisely, it is known that darker skin has larger amounts of melanin than pale skin, rather than more melanocytes

(Chapter 28). The melanin distribution also varies across skin types, as the protective melanin cap is superior in black skin in its arrangement over keratinocyte nuclei. So the genes controlling melanin production and deposition in the cell are potentially important etiologically (Boissy, 2003). Melanin is an effective sunblock with-two- to threefold differences in the black melanin pigment (eumelanin) potentially accounting for up to a hundredfold variation in sun sensitivity (Chapter 28).

Pheomelanin is another form of melanin which is yellow/red in color and is less photo-protective. Pheomelanin is produced particularly in people with red hair and freckles. Red hair is a risk factor for melanoma (Osterlind *et al.*, 1988), as is a reported tendency to sunburn. These phenotypic characteristics are thought to correlate with pheomelanin production in the skin.

The ratio of pheomelanin to eumelanin produced by the melanocyte is modulated by the melanocortin receptor MC1R (Suzuki, Cone *et al.*, 1996). *MC1R* variants differ in frequency across populations. They are common in some Northern European peoples, particularly of Celtic derivation; they are also seen in Asian people but are rare in black peoples (Harding *et al.*, 2000). Specific variants of the *MC1R* gene have been shown to be associated with red hair (Valverde *et al.*, 1995) and freckles (Bastiaens *et al.*, 2001). In an Australian population, four variants have been shown to be strongly predictive of red hair and fair skin (Arg151Cys, Asp84Glu, Arg160Trp and Asp295His) (referred to as R alleles) and three were less strongly associated (Val160Leu, Val92Met and Arg163Gln) (r alleles) (Duffy *et al.*, 2004). Arg142His and His260Pro alleles are reported as additional variants predictive of red hair in the Dutch population (Kennedy *et al.*, 2001). In epidemiological studies, inheritance of the ''red hair'' variants are risk factors for melanoma (Valverde *et al.*, 1996), which is not surprising given the strength of the association for the red hair phenotype. However, genetic analysis suggests that these variants may confer risk even in people with dark hair (Kennedy *et al.*, 2001; Palmer *et al.*, 2000). This needs to be evaluated in other studies as the possibility remains that this could represent incomplete recording of the true ''risk'' phenotype, as for instance, dark haired individuals often have a red beard.

Conclusions

This chapter has attempted to describe some of the initial studies investigating the joint effects of genes and environment for cancer, which is often termed ''gene–environment interaction''. The reason for preferring the more generic term ''joint effects of genes and environment'' is to avoid focusing on statistical interactions which imply the examination of a precise statistical hypothesis. The examples indicate that there has been some limited progress in these investigations. There are a number of complexities to these analyses, including the understanding of the critical biological processes and the knowledge and measurement of the exposures. To date, the statistical evaluation of the joint effects of genes and environment rely on simplistic analyses which have a limited basis in biology. More meaningful analyses are likely to require more sophisticated models which build on the biology and the likely impact of the exposure (Thomas, 2000).

REFERENCES

Ames, B.N. (1999). Micronutrient deficiencies. A major cause of DNA damage. *Ann N Y Acad Sci*, **889**, 87–106.

Antoniou, A.C., Pharoah, P.D., Mcmullan, G. *et al.* (2002). A comprehensive model for familial breast cancer incorporating *BRCA1*, *BRCA2* and other genes. *Br J Cancer*, **86**, 76–83.

Arlett, C.F. and Lehmann, A.R. (2004). Xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy: sun sensitivity, DNA repair defects and skin cancer. In Eeles, R., Easton, D.F., Ponder, B.A.J. and Eng, C., eds., *Genetic predisposition to cancer*. 2nd edn. London: Arnold.

Armstrong, B. K. (1988). Epidemiology of malignant melanoma: intermittent or total accumulated exposure to the sun? *J Dermatol Surg Oncol*, **14**, 835–49.

Barrett, J. H., Smith, G., Waxman, R. *et al.* (2003). Investigation of interaction between N-acetyltransferase 2 and heterocyclic amines as potential risk factors for colorectal cancer. *Carcinogenesis*, **24**, 275–82.

Bastiaens, M., Ter Huurne, J., Gruis, N. *et al.* (2001). The melanocortin-1-receptor gene is the major freckle gene. *Hum Mol Genet*, **10**, 1701–8.

Bishop, D. T., Demenais, F., Goldstein, A. M. *et al.* (2002). Geographical variation in the penetrance of *CDKN2A* mutations for melanoma. *J Natl Cancer Inst*, **94**, 894–903.

Boissy, R. E. (2003). Melanosome transfer to and translocation in the keratinocyte. *Exp Dermatol*, **12** Suppl 2, 5–12.

Brunet, J. S., Ghadirian, P., Rebbeck, T. R. *et al.* (1998). Effect of smoking on breast cancer in carriers of mutant *BRCA1* or *BRCA2* genes. *J Natl Cancer Inst*, **90**, 761–6.

Chen, J., Stampfer, M. J., Hough, H. L. *et al.* (1998). A prospective study of N-acetyltransferase genotype, red meat intake, and risk of colorectal cancer. *Cancer Res*, **58**, 3307–11.

Chuang, T. Y., Charles, J., Reizner, G. T. *et al.* (1999). Melanoma in Kauai, Hawaii, 1981–1990: the significance of in situ melanoma and the incidence trend. *Int J Dermatol*, **38**, 101–7.

Cunningham, J. M., Christensen, E. R., Tester, D. J. *et al.* (1998). Hypermethylation of the hMLH1 promoter in colon cancer with microsatellite instability. *Cancer Res*, **58**, 3455–60.

Duffy, D. L., Box, N. F., Chen, W. *et al.* (2004). Interactive effects of *MC1R* and *OCA2* on melanoma risk phenotypes. *Hum Mol Genet*, **13**, 447–61.

Eeles, R., Easton, D. F., Ponder, B. A. J. and Eng, C., eds. (2004). *Genetic predisposition to cancer*. London: Arnold.

Ewens, W. J. and Spielman, R. S. (1995). The transmission/disequilibrium test: history, subdivision, and admixture. *Am J Hum Genet*, **57**, 455–64.

Gauderman, W. J., Witte, J. S. and Thomas, D. C. (1999). Family-based association studies. *J Natl Cancer Inst Monogr*, 31–7.

Ghadirian, P., Lubinski, J., Lynch, H. *et al.* (2004). Smoking and the risk of breast cancer among carriers of *BRCA* mutations. *Int J Cancer*, **110**, 413–16.

Goldgar, D. E., Easton, D. F., Cannon-Albright, L. A. and Skolnick, M. H. (1994). Systematic population-based assessment of cancer risk in first-degree relatives of cancer probands. *JNCI*, **86**, 1600–8.

Goldstein, A. M. and Andrieu, N. (1999). Detection of interaction involving identified genes: available study designs. *J Natl Cancer Inst Monogr*, 49–54.

Haenszel, W. (1982). Migrant studies. In Schottenfeld, D. and Fraumeni, J. F., Jnr., eds., *Cancer epidemiology and prevention*. Philadelphia: W.B. Saunders Company.

Harding, R. M., Healy, E., Ray, A. J. *et al.* (2000). Evidence for variable selective pressures at MC1R. *Am J Hum Genet*, **66**, 1351–61.

Hemminki, K., Li, X. and Czene, K. (2004). Familial risk of cancer: data for clinical counseling and cancer genetics. *Int J Cancer*, **108**, 109–14.

Hopper, J. L. (2003). Commentary: Case-control-family designs: a paradigm for future epidemiology research? *Int J Epidemiol*, **32**, 48–50.

Hopper, J. L. and Carlin, J. B. (1992). Familial aggregation of a disease consequent upon correlation between relatives in a risk factor measured on a continuous scale. *Am J Epidemiol*, **136**, 1138–47.

Hudson, D. A. and Krige, J. E. J. (1993). Plantar melanoma in black South Africans. *British Journal of Surgery*, **80**, 992–4.

Jimbow, K. and Kukita, A. (1984). Unique location of melanoma, types of chemotherapy and immunotherapy, and survival of melanoma patients in Japan. In Mackie R., ed., *Clin Oncol*, **3**, 477–92.

Jones, W., Harman, C., Ng, A. and Shaw, J. (1999). The incidence of malignant melanoma in Auckland, New Zealand: highest rates in the world. *World J Surg*, **23**, 732–5.

Kadlubar, F. F., Butler, M. A., Kaderlik, K. R. *et al.* (1992). Polymorphisms for aromatic amine metabolism in humans: relevance for human carcinogenesis. *Environ Health Perspect*, **98**, 69–74.

Kampman, E., Slattery, M. L., Bigler, J. *et al.* (1999). Meat consumption, genetic susceptibility, and colon cancer risk: a United States multicenter case-control study. *Cancer Epidemiol Biomarkers Prev*, **8**, 15–24.

Kennedy, C., Ter Huurne, J., Berkhout, M. *et al.* (2001). Melanocortin 1 receptor (*MC1R*) gene variants are associated with an increased risk for cutaneous melanoma which is largely independent of skin type and hair color. *J Invest Dermatol*, **117**, 294–300.

Kraemer, K. H. and Slor, H. (1985). Xeroderma pigmentosum. *Clin Dermatol*, **3**, 33–69.

Norat, T., Bingham, S., Ferrari, P. *et al.* (2005). Meat, fish, and colorectal cancer risk: the European Prospective

Investigation into cancer and nutrition. *J Natl Cancer Inst*, **97**, 906–16.

Osterlind, A., Tucker, M.A., Hou-Jensen, K. *et al.* (1988). The Danish case-control study of cutaneous malignant melanoma. I. Importance of host factors. *Int J Cancer*, **42**, 200–6.

Palmer, J.S., Duffy, D.L., Box, N.F. *et al.* (2000). Melanocortin-1 receptor polymorphisms and risk of melanoma: is the association explained solely by pigmentation phenotype? *Am J Hum Genet*, **66**, 176–86.

Parkin, D., Whelan, S., Ferlay, J. *et al.* (1997). Cancer incidence in five continents. *IARC Scientific Publications No. 143. II*. Lyon: International Agency for Research on Cancer.

Peto, J., Collins, N., Barfoot, R. *et al.* (1999). Prevalence of *BRCA1* and *BRCA2* gene mutations in patients with early-onset breast cancer. *J Natl Cancer Inst*, **91**, 943–9.

Potter, J.D. and Hunter, D. (2002). Colorectal cancer. In Adami, H.-O., Hunter, D. and Trichopoulos, D., eds., *Textbook of cancer epidemiology*. Oxford: Oxford University Press.

Potter, J.D., Slattery, M.L., Bostick, R.M. and Gapstur, S.M. (1993). Colon cancer: a review of the epidemiology. *Epidemiol Rev*, **15**, 499–545.

Rivers, J.K. (2004). Is there more than one road to melanoma? *Lancet*, **363**, 728–30.

Roberts-Thomson, I.C., Ryan, P., Khoo, K.K. *et al.* (1996). Diet, acetylator phenotype, and risk of colorectal neoplasia. *Lancet*, **347**, 1372–4.

Slattery, M.L., Curtin, K., Anderson, K. *et al.* (2000). Associations between cigarette smoking, lifestyle factors, and microsatellite instability in colon tumors. *J Natl Cancer Inst*, **92**, 1831–6.

Slattery, M.L., Potter, J.D., Samowitz, W. *et al.* (1998). *NAT2*, *GSTM-1*, cigarette smoking, and risk of colon cancer. *Cancer Epidemiol Biomarkers Prev*, **7**, 1079–84.

Stevens, N.G., Liff, J.M. and Weiss, N.S. (1990). Plantar melanoma: is the incidence of melanoma of the sole of the foot really higher in Blacks than Whites? *International Journal of Cancer*, **45**, 691–3.

Suzuki, I., Cone, R.D., Im, S. *et al.* (1996). Binding of melanotropic hormones to the melanocortin receptor *MC1R* on human melanocytes stimulates proliferation and melanogenesis. *Endocrinology*, **137**, 1627–33.

Thomas, D.C. (2000). Genetic epidemiology with a capital ''E''. *Genet Epidemiol*, **19**, 289–300.

Valverde, P., Healy, E., Jackson, I. *et al.* (1995). Variants of the melanocyte stimulating hormone receptor gene are associated with red hair and fair skin in humans. *Nature Genetics*, **11**, 328–30.

Valverde, P., Healy, E., Sikkink, S. *et al.* (1996). The Asp84Glu variant of the melanocortin 1 receptor (*MC1R*) is associated with melanoma. *Hum Mol Genet*, **5**, 1663–6.

Vogelstein, B., Fearon, E.R., Kern, S.E. *et al.* (1989). Allelotype of colorectal carcinomas. *Science*, **244**, 207–11.

Welfare, M.R., Cooper, J., Bassendine, M.F. and Daly, A.K. (1997). Relationship between acetylator status, smoking, and diet and colorectal cancer risk in the north-east of England. *Carcinogenesis*, **18**, 1351–4.

15

The polygenic basis of breast cancer

Paul D. P. Pharoah and Bruce A. J. Ponder

Introduction

A major challenge for molecular and epidemiological science is to unravel the molecular genetic basis of chronic disease. The past 30 years have seen considerable progress in understanding the molecular genetics of diseases that are inherited according to Mendelian rules, that is, those in which mutations in a single gene have a large effect on disease risk (see Chapter 10). However, little is known about complex disease, which results from the combined effects of many genes, that is, diseases which show polygenic inheritance. Rapid advances in our understanding of human genome architecture together with technological developments may help us meet this challenge successfully. This will bring new insights into disease etiology, which, in turn, will help in the development of new methods for disease prevention and treatment. As a result of these advances, it may also become possible to target interventions to individuals at greatest risk of disease. In this chapter we will review the evidence for the polygenic model of breast cancer susceptibility and discuss the implications of the model for disease prevention in the population. We will contrast this with the potential impact of preventive interventions targeted at women with single gene disorders who are at high risk of disease.

Genetic models of breast cancer susceptibility

It is likely that the inheritance of most common cancers is polygenic. Breast cancer, like other common cancers, exhibits some degree of familial clustering, with disease being approximately twice as common in first-degree relatives of cases (Amundadottir *et al.*, 2004; Collaborative Group on Hormonal Factors in Breast Cancer, 2001). The higher rate of most cancers in the monozygotic twins of cases than in dizygotic twins or siblings suggests that most of the familial clustering is the result of genetic variation rather than lifestyle or environmental factors (Lichtenstein *et al.*, 2000; Peto and Mack, 2000). Further evidence for the relative importance of genetic factors comes from the observation that more distant relatives of a case (i.e. those beyond the nuclear family) are also at increased risk of disease even though they would be expected to share environmental or lifestyle factors to a lesser degree (Amundadottir *et al.*, 2004). Furthermore, the magnitude of the risks in distant relatives are close to those predicted by simple genetic models, such as a dominant model or an additive polygenic model.

Some of the familial clustering of breast cancer occurs as part of specific inherited breast cancer syndromes, where disease results from single genes conferring a high risk. Several genes associated

with these syndromes have been identified, including *BRCA1*, *BRCA2*, *PTEN* and *TP53*. However, the susceptibility alleles in these genes are rare in the population and they account for a small minority of the inherited component of cancer. Highly penetrant variants in the breast cancer susceptibility genes *BRCA1* and *BRCA2* account for less than 20% of the genetic risk of breast cancer, with other rarer high penetrance genes such as *TP53*, *ATM* and *PTEN* accounting for less than 5% (Easton, 1999). Other *BRCA1/2*-like genes are unlikely to exist, as the majority of multiple case families can be accounted for by *BRCA1* or *BRCA2* (Ford *et al.*, 1998) and, despite extensive research efforts, attempts to identify similar highly penetrant cancer susceptibility genes have failed.

Mathematical modeling of the familial aggregation of breast cancer in the population can provide important clues about the range of genetic models that best account for the familial aggregation of breast cancer not due to the high-penetrance *BRCA* genes. Two such studies have been published. The first study used data from a series of 856 breast cancer cases from Australia that were diagnosed in women under the age of 40 and had been tested for mutations in *BRCA1* and *BRCA2* (Cui *et al.*, 2001). The genetic model that fitted these data the best was that of a single recessive allele which conferred a high disease risk. Another study analyzed the occurrence of breast cancer in the relatives of patients in the Anglian Breast Cancer Study (2000), a population-based series of ~1500 cases, all of whom were screened for mutations in *BRCA1* and *BRCA2* (Antoniou *et al.*, 2001). Two genetic models were found to fit the data well. The model best describing these data was a polygenic model, in which susceptibility to breast cancer is conferred by a large number of alleles. The risk associated with any individual allele is small, but the effects are multiplicative so that a woman with several susceptibility alleles is at high risk. The second model was that of a single common recessive allele (frequency 0.24). The hypothetical recessive risk allele was estimated to confer a relative risk of 21 for rare homozygotes, compared with common homozygotes and heterozygotes, corresponding to a moderately high penetrance of 42%. The polygenic and recessive models were also applied to a series of multiple case families not due to *BRCA* mutations (Antoniou *et al.*, 2001). The polygenic model fitted these data well, but the fit of the recessive model was not so good. Furthermore, a recent large meta-analysis found that the familial breast cancer risk to siblings is similar to that to mothers, suggesting that any recessive component for the excess familial risk is at best small (Collaborative Group on Hormonal Factors in Breast Cancer, 2001).

The evidence for the polygenic model of breast cancer susceptibility is persuasive and it is likely to be an appropriate model for many common cancers and other diseases. However, the number of risk alleles and their properties (allele frequencies and risks conferred) are not known. Indeed, a wide variety of possibilities are consistent with a polygenic model for cancer susceptibility, ranging from a handful of alleles of moderate risk to a large number of alleles, each conferring a slight increase in risk across a wide range of allele frequencies, or a combination of the two. Despite this uncertainty, there are some general implications that can be drawn that do not depend on which of these possibilities turns out to be correct.

Implications of the polygenic model

A general polygenic model has implications for the distribution of breast cancer risk in the population. These are illustrated in the following paragraphs and in Figure 15.1. On a relative scale, the average population disease risk is, by definition, unity. Let us assume that one dominant susceptibility allele has been identified, and the risk allele frequency is 0.29 so that half the population carry at least one copy of the high risk allele. If ''high-risk'' individuals have a breast cancer relative risk (RR) of 3 compared to ''low-risk'' individuals, the high-risk group will have a RR of 1.5 and the low-risk group will have a RR of 0.5 ($1.5 = 3 \times 0.5$), giving a population average RR of one. Absolute risks of

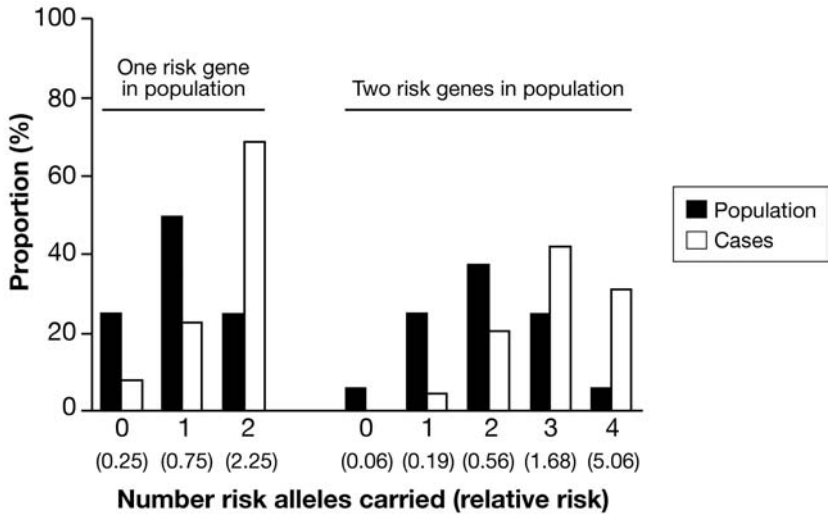

Figure 15.1 Proportion of population and cases by number of risk alleles carried for (*A*) a single codominant gene (*B*) two codominant genes. Each additional risk allele carried results in a threefold increase in risk.

breast cancer will depend on underlying incidence rates. Using rates typical of northern Europe and the United States, the absolute risk of breast cancer by age 70 in the two risk groups will be 2.9% and 8.4%, corresponding to an average population risk of 5.7% (Pharoah and Mackay, 1998). It then follows that 75% of all breast cancer cases will occur in high-risk women and 25% in low-risk women. An intervention targeted to ''high-risk'' women thus has the potential to reduce breast cancer morbidity by a maximum of 75%. Now assume that there are two risk alleles in two genes each with allele frequency 0.5 and that these act in a multiplicative, co-dominant manner. There will now be three risk groups in the population (Figure 15.1): one quarter of the population will carry no high-risk alleles and will have a RR of 0.25 (0.5×0.5); half the population will carry one high-risk allele (RR of 0.75); and one quarter of the population will carry two high-risk alleles (RR of 2.25). The breast cancer risks by age 70 in these three groups are 1.5%, 4.3% and 12.3% and they account for 8%, 23% and 69% of all breast cancer cases respectively. Figure 15.1 also shows the effect of two co-dominant genes with five risk groups. As the number of susceptibility genes increases, the number of risk groups will increase and disease risk in the population tends towards a continuous distribution.

Under the polygenic model, the (continuous) distribution of RR in the population is predicted to be log-normal; that is, the logarithm of RR for all individuals in the population will follow a normal distribution (Figure 15.2). A normal distribution is defined by its mean value and its standard deviation. The standard deviation of the log-normal distribution of the polygenic risk for breast cancer was estimated to be 1.2 from a population-based study (Antoniou *et al.*, 2001). Once the standard deviation is defined, the mean of the distribution can be set so that the arithmetical average risk (RR) is equal to 1. The standard deviation describes the variation in risks that can be defined within the population, and thus is the key indicator of the power to discriminate individuals in groups at low or high risk.

It can be shown that the distribution of (initial) risk among cases of the disease is also log-normal and has a simple relationship to the distribution of risk in the population. The standard deviation

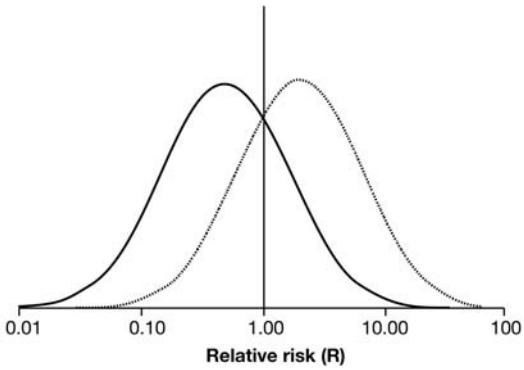

Figure 15.2 Distribution of polygenic risk in the population and in cases. Relative risks are shown on a log scale, while the arithmetical average risk for the entire population has been set at 1.0. The risk distribution in individuals who will develop breast cancer (cases) is shifted to the right. The standard deviation describes the spread of risk between high and low values within the population, and thus the potential to discriminate different levels in different individuals.

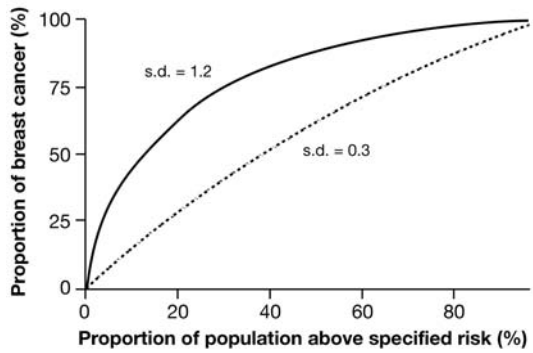

Figure 15.3 The proportion of cases accounted for by a given proportion of the population above a specified risk according to the standard deviation of underlying risk distribution. Thus the 12% of the population with the highest risk account for half of all cases if SD = 1.2, whereas for SD = 0.3, 38% of population account for half the cases.

of the log-normal distribution in cases is the same as in the population, but the average risk is higher (Figure 15.2). A minor complication occurs because at older ages the distribution of risk in both the general population and among cases changes as higher-risk individuals are more likely to have been ''eliminated''. Figure 15.2 shows the distribution of genetic risk in the population for a standard deviation of 1.2. The area under the curve gives the proportion of the population in any risk group. The risk to the highest quintile of the distribution is 40-fold higher than that of the lowest quintile. The figure also shows the risk distribution in cases according to their initial level of risk; that is, the risk distribution of women in the population who will subsequently develop breast cancer.

The proportion of the population that have a risk greater than a given level, and the proportion of cases that will occur within this high-risk subgroup, provides more useful information. These figures are obtained from the area under the population and case curves in Figure 15.2 to the right of any given risk cut-off. Thus, for a standard deviation of 1.2, half the population have a RR of 0.46 or higher and this half of the population account for 88% of all breast cancer cases. It can also be shown that the 12% of the population at highest risk account for 50% of cases. Another way of depicting these data is to plot the proportion of cases that occur in women above a given level of risk against the proportion of the population above that level of risk as shown in Figure 15.3.

In practice, the estimated power of the genetic risk distribution is an upper limit, because some disease-associated genes may prove difficult to detect, and this will reduce the width of the distribution and hence the predictive value. We therefore recalculated the model assuming that genes responsible for half the variation in genetic risk could be identified. The risk to the highest quintile of the distribution is now 12-fold higher than that of the lowest quintile. The results are shown in Figure 15.3.

The assumption that the putative polygenes act in a multiplicative manner may not be correct. Risch has argued, on the basis of the ratio of risks to monozygotic and dizygotic twins of cancer cases, that an additive model provides the best fit for most common cancers including breast cancer

(Risch, 2001). The effect of an additive model would be to reduce the standard deviation of the risk distribution from 1.2 to 1.05, and hence slightly reduce the predictive power of genetic testing. In contrast, a recent analysis of twin data reported by Peto and Mack found a very high incidence in monozygotic twins of patients, more consistent with a muliplicative model, and concluded that a high proportion, and perhaps the majority, of breast cancers arise in a susceptible minority of women (Peto and Mack, 2000). If this were true, the discriminatory power could be substantially improved.

Implications of the recessive model

Under the recessive model there would be a single risk allele with frequency 0.24. Thus women with the at risk genotype (RR = 9.8) would comprise 6% of the population and account for 56% of cases. The remaining 94% of the population (RR = 0.46) would account for 44% of cases. However, in view of the failure of genetic linkage studies to identify further breast cancer susceptibility genes, such a model seems less plausible than a polygenic model of inheritance (Risch, 2001; Vaittinen and Hemminki, 2000).

Public health implications

Disease prevention in those with single disease susceptibility alleles

The potential benefits of disease prevention in individuals who carry a high penetrance gene are clear, and provide a rationale for genetic testing in a family cancer clinic. What then might be the impact of the identification of genetic risk factors for disease, in terms of disease prevention at the level of the population? Several authors have pointed out that individual susceptibility genes are unlikely to contribute much to disease prevention (Vineis *et al.*, 2001). For example, consider a highly penetrant gene mutation that is rare in the population (0.2%). Suppose that the mutation confers a tenfold increase in risk in carriers and the lifetime risk of disease in a carrier is 50%. The mutation would be present in 2% of all cases. If we have an intervention that reduces risk by 40% the absolute risk reduction in carriers is 20% (40% of 50%). Thus for every 5 (100/20) carriers treated, we would prevent one case. This is the ''number needed to treat'' (NNT). However, if we identified carriers by testing (or ''screening'') the population, we would need to screen 2500 individuals (5/0.002) to prevent one case. This is the number needed to screen (NNS). A population screening programme to detect and treat carriers would reduce total disease burden by 0.8% if uptake of testing and treatment were complete.

Let us now consider a more common, low-penetrance genetic variant, which carries a twofold increase in disease risk and a lifetime risk of disease of 10%, and is present in 5% of the population. The variant would be present in 9.5% of all cases. An intervention that reduces risk by 40% (absolute risk reduction 4%) will have an NNS of 500, and, at best, could reduce total disease burden by 3.8%.

Disease prevention under the polygenic model

The possibility that genetic susceptibility to breast cancer is due to several loci, each conferring a modest independent risk, seems reasonable. In practice, the number of loci involved will be finite, but once there are more than 4–5 loci the distribution of risk will be similar to the polygenic model except at the extreme tails. A key aspect of the model is the standard deviation (σ) of the polygenic risk distribution, because this determines the power of the distribution to discriminate high- and low-risk individuals. The estimate of the standard deviation obtained by Antoniou *et al.* is a property of the segregation analysis model (Antoniou *et al.*, 2001) and it is also close to that predicted by other studies of familial risk. The RR of disease in monozygotic twins ($\lambda_{monozygotic}$) and siblings ($\lambda_{sibling}$) are related to each other and to the

predicted standard deviation (σ) of the polygenic log-normal risk distribution by the equation:

$$\lambda_{\text{monozygotic}} = \lambda^2_{\text{sibling}} = e^{\sigma^2}$$

Assuming λ_{sibling} to be equal to 2, as estimated by many observational epidemiologic studies, this equation solves to predict a standard deviation of 1.2. The familial RRs for many other common cancers are also around 2 to 3 (Amundadottir *et al.*, 2004; Goldgar *et al.*, 1994), which suggests that the distribution of risk for these cancers will be similar to what we have observed for breast cancer. Thus, the potential benefits of a targeted high-risk approach to disease prevention are also likely to be similar.

The practical use of risk information should be considered in two contexts; that of the individual, and that of the population as defined by Rose (Rose, 1985). In both cases, our analysis suggests that a ''risk profile'' that is based on the combination of known genotype and other risk factors is likely to provide risk discrimination which will be of practical value in healthcare terms. Whether genetic testing in whole populations would be socially or economically acceptable remains unknown, and is likely to depend on whether useful action can be seen to result. But it does seem clear that the use of combinations of risk factors is potentially able to overcome many of the limitations of single-risk factors, which have been the cause of scepticism about the practical utility of molecular genotyping for common, low-risk genes (Holtzman and Marteau, 2000; Vineis *et al.*, 2001).

For example, in respect of individual risk, a single gene which conferred a RR of breast cancer of 1.5-fold – the size of effect that seems plausible from reported studies (Dunning *et al.*, 1999) – would increase the risk of breast cancer to an individual from the UK population average of 5.7% by age 70 to around 8.5%. On the other hand, inspection of Figure 15.2 suggests that a genotypic risk profile might identify one woman in 30 who has a risk by age 70 of 20% or more. Little is known about how individuals will perceive and respond to such risks; but the discriminatory power of the polygenic risk profile is clear.

At the population level, the effects are even more striking. According to the genetic model, 12% of the population have a risk of breast cancer of 1 in 10 or more by age 70; and half the total breast cancer incidence falls within that 12% of the population. Different cut-offs can be chosen to give the best combination of high risk and proportion of total breast cancer incidence that is included within the high risk group, to suit the purpose in hand. A single genotypic marker would, by contrast, provide far weaker discrimination: for example, a dominant predisposing allele with frequency 10% and RR 1.5 would result in 26% of cancer incidence occurring in the 19% of the population who carried at least one allele; very poor enrichment in terms of targeting interventions. An important feature of the high-risk groups defined by the model is that most of the individuals within them will be at risk because of the combined effect of several predisposing alleles. This implies that interventions that are based on specific mechanisms of predisposition will individually deal with only a proportion of the excess cancer risk; and that except for predisposing genes with major effects, generic interventions are more likely to be appropriate.

Risk profiles may also be used to define low-risk groups. Thus, in Figure 15.2, only 12% of breast cancer incidence falls within the 50% of women at lowest risk. Exclusion of the low-risk groups from interventions, if it were socially acceptable, might be very cost-effective. For example, screening of the whole population by mammography should reduce breast cancer mortality by approximately 30% (Kerlikowske *et al.*, 1995). If mammography were offered only to the half of the population in the highest risk group by the genetic profile, total breast cancer mortality would still be reduced by 26%; a ''loss'' of only 4%. There would be additional benefits, since the benefit:harm ratio is likely to be improved by targeting to the high-risk group. (These arguments assume that the efficacy of any intervention is

independent of genotype: if that is not the case, the dividend from genotyping may be greater or smaller depending on whether the cancers in high-risk individuals are more or less responsive to the intervention).

These arguments and examples assume that all of the genetic factors that contribute to the estimated risk distribution can be identified and typed. In practice, this goal is some way off (Pharoah *et al.*, 2004). Nevertheless, the figures in the Table suggest that even if only half the risk factors were typed, useful discrimination of risk might be possible.

Finding low-penetrance breast cancer alleles

The genetic association study has been promoted as the most efficient method to identify common low-penetrance disease suceptibility alleles (Risch and Merikangas, 1996). This design has been used as a method to map cancer susceptibility alleles for over four decades, but the advent of modern molecular genetics has seen a dramatic increase in the use of this type of study in the past five years (Pharoah *et al.*, 2004). These efforts have, perhaps, been most notable for their few successes. However, most studies carried out so far have been based on the candidate gene/candidate polymorphism approach, which has been limited by the availability of plausible candidates. In addition, many studies have lacked statistical power to detect moderate risks because limitations in molecular genetic technologies have limited sample sizes. In the future, improved knowledge of the molecular basis of carcinogenesis will enable better selection of candidate genes, and improvements in study design, notably the use of much larger sample sizes through multicenter collaborations, may herald a new era of susceptibility gene discovery. An example of this type of collaboration is the Consortium of Cohorts formed by the National Cancer Institute to address the need for large-scale collaborations for study of gene–gene and gene–environment interactions in the etiology of cancer with more than 20 cohorts participating (www.epi.grants.cancer.gov/Consortia/cohort.html). Advances in our understanding of the nature of human genetic variation coupled to new genotyping technologies raise the hope that empirical whole-genome approaches will bring similar successes to those achieved by empirical family-based linkage studies. Empirical association studies to scan 60% of the genome for breast cancer susceptibility genes are currently in progress, and similar studies of other cancers are likely in the near future.

The association studies design relies on the ''common disease: common variant'' hypothesis. It is, however, equally possible that much of the variation in cancer risk is due to rarer alleles. Indeed, virtually all susceptibility alleles identified to date have frequencies of less than 1%. These include both high-penetrance mutations, but also low-penetrance variants in *ATM* and *CHEK2* that predispose to breast cancer. Although the identification of rare variants that confer modest risks is possible, the problems are formidable. Firstly, very large sample sizes are required: a recent association study to confirm the breast cancer risk associated with the 1100del*c* variant of *CHEK2* in breast cancer included close to 10 000 cases and 10 000 controls. Secondly, much less is known about the occurrence of rare variants across the genome as current efforts to identify genetic variation have concentrated on common polymorphisms identifiable by resequencing small numbers of individuals. Identifying rare variants will require resequencing much larger numbers of individuals, perhaps concentrating on individuals from high risk families where the frequencies may be higher. Furthermore, because the numbers of rare variants will be much larger, the problems of multiple testing will be much greater. Finally, the frequencies of such alleles are likely to vary between populations.

The relative importance of common and rare variants in cancer susceptibility cannot be inferred from existing data and remains a major uncertainty

overhanging future studies. The identification of common alleles is more tractable and, because they would explain a greater fraction of disease burden, is of more direct public health relevance. It is therefore logical to concentrate on these. However, alleles are likely to be rare if there is any degree of selection against the variant, for example if homozygotes are non-viable. If most cancer susceptibility is related to fundamental processes of cellular control, rare alleles may turn out to be the more important component.

Conclusion

There is good evidence to support the polygenic model for breast cancer susceptibility and it seems reasonably likely that the polygenic model will apply to the other common cancers. If all the disease associated alleles that contribute to breast cancer susceptibility were known, the potential benefits of targeted preventive interventions would be clear, at both the level of the individual and when considering disease burdens in populations. A major challenge for molecular genetics and genetic epidemiology is to identify some or all of these alleles.

REFERENCES

Amundadottir, L. T., Thorvaldsson, S., Gudbjartsson, D. F. *et al.* (2004). Cancer as a complex phenotype: pattern of cancer distribution within and beyond the nuclear family. *PLoS Med*, **1**, e65.

Anglian Breast Cancer Study Group. (2000). Prevalence and penetrance of *BRCA1* and *BRCA2* in a population based series of breast cancer cases. *Br J Cancer*, **83**, 1301–8.

Antoniou, A. C., Pharoah, P. D. P., McMullen, G. *et al.* (2001). Evidence for further breast cancer susceptibility genes in addition to *BRCA1* and *BRCA2* in a population based study. *Genetic Epidemiology*, **21**, 1–18.

Collaborative Group on Hormonal Factors in Breast Cancer. (2001). Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet*, **358**, 1389–99.

Cui, J., Antoniou, A. C., Dite, G. S. *et al.* (2001). After *BRCA1* and *BRCA2* – what next? Multifactorial segregation analyses of three-generation, population-based Australian families affected by female breast cancer. *Am J Hum Genet*, **68**, 420–31.

Dunning, A. M., Healey, C. S., Pharoah, P. D. P. *et al.* (1999). A systematic review of genetic polymorphisms and breast cancer risk. *Cancer Epidemiology Biomarkers and Prevention*, **8**, 843–54.

Easton, D. F. (1999). How many more breast cancer predisposition genes are there. *Breast Cancer Research*, **1**, 14–17.

Ford, D., Easton, D. F., Stratton, M. *et al.* (1998). Genetic heterogeneity and penetrance analysis of the *BRCA1* and *BRCA2* genes in breast cancer families. *American Journal of Human Genetics*, **62**, 676–89.

Goldgar, D. E., Easton, D. F., Cannon-Albright, L. A. and Skolnick, M. H. (1994). A systematic population based assessment of cancer risk in first degree relatives of cancer probands. *Journal of the National Cancer Institute*, **86**, 1600–8.

Holtzman, N. A. and Marteau, T. M. (2000). Will genetics revolutionize medicine? *N Engl J Med*, **343**, 141–4.

Kerlikowske, K., Grady, D., Rubin, S. M., Sandrock, C. and Ernster, V. L. (1995). Efficacy of screening mammography: a meta-analysis. *Journal of the American Medical Association*, **273**, 149–54.

Lichtenstein, P., Holm, N. V., Verkasalo, P. K. *et al.* (2000). Environmental and heritable factors in the causation of cancer – analyses of cohorts of twins from Sweden, Denmark and Finland. *N Engl J Med*, **343**, 78–85.

Peto, J. and Mack, T. M. (2000). High constant incidence in twins and other relatives of women with breast cancer. *Nat Genet*, **26**, 411–14.

Pharoah, P. D. P., Dunning, A. M., Ponder, B. A. J. and Easton, D. F. (2004). Association studies for finding cancer-susceptibility genetic variants. *Nature Reviews Cancer*, **4**, 850–60.

Pharoah, P. D. P. and Mackay, J. (1998). Absolute risk of breast cancer in women at increased risk: a more useful clinical measure than relative risk? *The Breast*, **7**, 255–9.

Risch, N. (2001). The genetic epidemiology of cancer: interpreting family and twin studies and their

implications for molecular genetic approaches. *Cancer Epidemiol Biomarkers Prev*, **10**, 733–41.

Risch, N. and Merikangas, K. (1996). The future of genetic studies of complex diseases. *Science*, **273**, 1516–17.

Rose, G. (1985). Sick individuals and sick populations. *International Journal of Epidemiology*, **14**, 32–8.

Vaittinen, P. and Hemminki, K. (2000). Risk factors and age-incidence relationships for contralateral breast cancer. *Int J Cancer*, **88**, 998–1002.

Vineis, P., Schulte, P. and McMichael, A. J. (2001). Misconceptions about the use of genetic tests in populations. *Lancet*, **357**, 709–12.

16

TP53: A master gene in normal and tumor suppression

Pierre Hainaut

Introduction

Since its discovery over 25 years ago, the *TP53* gene is one of the ''stars'' of molecular cancer research. The p53 protein acts as an all-round regulator of many interconnected functions associated with cell cycle regulation, apoptosis, DNA repair, differentiation, senescence and development. Activation of p53 prevents DNA replication and cell proliferation when cells are subjected to stress that may perturb genetic or genomic integrity. Thus, *TP53* acts as a ''master suppressor gene'' by exerting simultaneous, many-fingered control of several components of the molecular mechanisms of carcinogenesis. Mutations in *TP53* result in loss of these suppressor functions. In some instances, it has been suggested that mutations may also exert gain-of-function effects that may explain the persistence of p53 mutant protein in cancer cells. *TP53* is emerging as an important target for improving cancer detection, prognosis and treatment. However, forms of mutant p53 differ from each other and this may affect cancer development in an organ, tissue and context-specific manner. Addressing this diversity is essential for developing cancer management strategies using p53 as a target.

Cancer progression is characterized by acquisition of multiple genetic and epigenetic alterations in genes involved in interrelated processes controlling cell cycle, apoptosis, differentiation, replicative senescence, cell motility and migratory capacity (Hanahan and Weinberg, 2000). There are many ways in which cells develop defects in these processes, often in a cell-type, tissue- and context-specific manner. However, a small number of genes are commonly altered in many different cancers, irrespective of their histology or site. These genes can be seen as ''bottlenecks'' or obligate crossing points on the path to cancer. The *TP53* tumor suppressor gene, encoding the p53 protein, is the prototype of such a ''master tumour gene''. With a mutation prevalence of 33% in all cancer lesions and over 50% in invasive cancers, *TP53* is the most frequently mutated cancer gene (Olivier *et al.*, 2002). However, mutations are only the most visible part of *TP53* involvement in cancer. Many other mechanisms contribute to inactivate p53 function in cancer cells, such as loss of alleles, sequestration by cellular or viral proteins, accelerated protein degradation, modification of protein shape, impairment of the activation process, and alteration of subcellular distribution (Hainaut and Hollstein, 2000). On the basis of our current knowledge of p53 function, cells which retain a perfectly functional p53 pathway are very unlikely ever to evolve into full-blown cancer cells. On the other hand, inactivation of p53 is not by itself sufficient for cells to cross the thin barrier between normality and malignancy.

After over 25 years of research and 30 000 publications, studies on *TP53* have had a major

impact on our understanding of the molecular biology of cancer. The challenge for the years to come is to turn this knowledge into advances in cancer prevention, detection, prognosis and therapy. In this chapter, I will discuss the main characteristics of *TP53*, its regulation, the consequences of its inactivation in cancer, and provide a perspective on exploitation of this knowledge both in molecular pathology and in the management of cancer.

Characteristics

Gene structure

The *TP53* gene (OMIM:191170, synonym: P53) is located on chromosome 17p13.1 and encodes an ubiquitous phosphoprotein of molecular mass 51–53000. The gene spans 20 kb, with a non-coding exon 1 and a very long first intron of 10 kb. The coding sequence contains five regions showing a high degree of conservation in vertebrates, predominantly in exons 2, 5, 6, 7 and 8, but the sequences found in invertebrates show only distant resemblance to mammalian *TP53* (May and May, 1999). Over 15 polymorphisms are identified in the human population, with allele frequencies that vary with ethnic origin (see list at www-p53.iarc.fr). One of them affects the coding sequence at codon 72, specifying either an arginine or a proline. The *Arg* allele is the most common in the western population (allele frequencies ranging from 0.6 to 0.8) but the prevalence of the *Pro* allele seems to increase according to a North–South gradient, so that the Pro allele is the most frequent one near the equator and in indigenous populations of the southern hemisphere (Beckman *et al.*, 1994). There is evidence that this polymorphism may have a functional impact on cancer susceptibility and response to therapy (Bergamaschi *et al.*, 2003). Interpreting the role of this polymorphism is however complicated by the fact that there are two intronic polymorphisms, located in introns 2 and 3, and in linkage disequilibrium with the codon 72 polymorphism. Further studies are needed to assess the specific role of each of these variations and to identify the structure of alleles that may carry a greater risk.

TP53 gene family

Two genes related to *TP53* have been identified on chromosome 1p36 (*TP73*, OMIM 601990) and on chromosome 3p28 (*TP63*, OMIM 603273). They both encode proteins with high homology to p53 in terms of overall structure. All three family members are sequence-specific transcription factors, with an acidic, N-terminal region, a central region that contains the DNA-binding domain, and a C-terminal region which contains, among other motifs, an oligomerization domain (Figure 16.1). They share the capacity to regulate the transcription of a common set of target genes (see below), suggesting the existence of some redundancy in their functions. Whereas *TP53* is ubiquitously expressed, the expression of *TP73* and *TP63* is tissue-specific. Experimental inactivation of these three genes by homologous recombination in mice results in very distinct phenotypes. *TP53*-deficient mice do not show major defects after birth, although a fraction of p53-deficient embryos display a lethal defect in neural tube closure, resulting in exencephaly. However, they are prone to the early formation of multiple, spontaneous cancers (Dumble *et al.*, 2003). In contrast, *TP73*-deficient mice do not show any evidence of enhanced susceptibility to cancer but have a complex pattern of epithelial, neurosensory and inflammatory defects. The most drastic phenotype is in *TP63*-deficient mice, which shows limb and cranial malformations, absence of prostatic and mammary gland buds, and inability to develop multilayered, squamous epithelia (Koster and Roop, 2004; Yang and McKeon, 2000). It is now emerging that the p63 protein is an essential regulator of the proliferation and differentiation of keratinocyte progenitors. Neither *TP73* nor *TP63* are mutated at a significant frequency in cancer (Blandino and Dobbelstein, 2004).

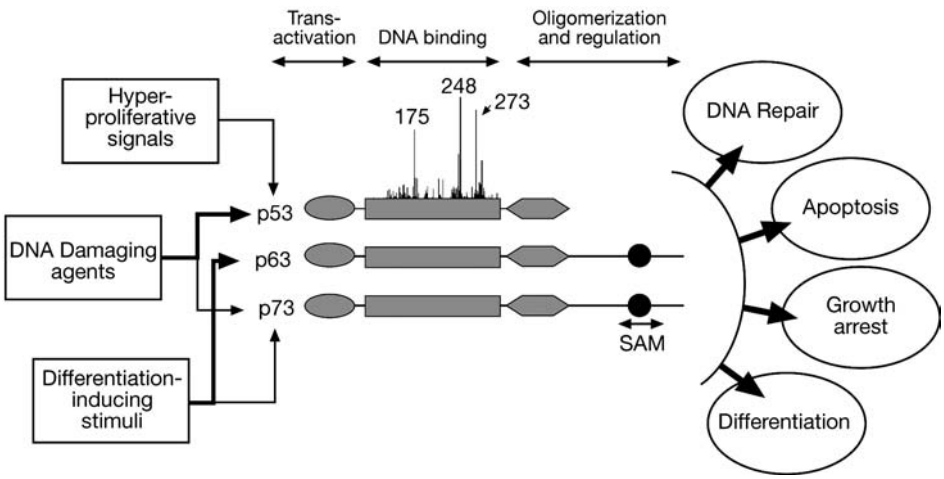

Figure 16.1 Schematic diagram of p53 showing the domain structure, major regulatory influences and effector outputs. The N-terminal transcriptional activation domain, central DNA binding domain and C-terminial oligomerization and regulatory domains, including the sterile alpha motif (SAM) protein interaction domain, are shown. Mutation ''hotspots'' at codons 175, 248 and 273 are indicated.

However, expression of these two genes is often deregulated, and *TP63* appears to be amplified and overexpressed in many squamous cell carcinomas of the upper airways and digestive tract (Massion *et al.*, 2003; Taniere *et al.*, 2001).

It has long been considered that *TP53* was essentially expressed as a single mRNA species. However, recent studies have shown that, similar to p63 and p73, p53 can be expressed as several isoforms resulting from alternative splicing or differential promoter usage that generate the production of multiple isoforms with variable N- or C-terminal regions (Murray-Zmijeswski *et al.*, 2006). The DeltaNp53 isoforms lack most of the N-terminal transactivation domain and may primarily act to block transactivation by its full-length counterpart, thus regulating p53 function (Courtois *et al.*, 2002). However, the exact biological role of p53 isoforms remains to be assessed.

p53 protein

The p53 protein contains 393 residues and is structured as a hydrophobic, central core (residues 110–296, encoded by exons 5 to 8), flanked by an acidic N-terminus and a basic C-terminus (Figure 16.1). The N-terminus contains two complementary transcriptional activation domains, with a major one at residues 1–42 and a minor one at residues 55–75, specifically involved in the regulation of several pro-apoptotic genes (Venot *et al.*, 1998). The central core is made up of a scaffold of two β-sheets supporting a set of flexible loops and helices stabilized by the binding of an atom of zinc. These loops and helices make direct contact with DNA sequences containing inverted repeats with the motif RRRC(A/T) (where R represents a purine base) (Cho *et al.*, 1994). The C-terminus contains the main nuclear localization signals and oligomerization domains (residues 325–366). The active form of the protein is a tetramer (in fact, a pair of dimers) (McLure and Lee, 1998). The extreme C-terminus has multiple regulatory functions and exerts a negative control on sequence-specific DNA binding activities. Both N- and C-terminal regions contain multiple phosphorylation sites, while none has been identified so far in the central core.

Regulation and function

Regulation of expression

The mechanisms of regulation of *TP53* transcription have been less studied than most other aspects of p53 protein function. The *TP53* gene does not contain a conventional TATA box but is under control of several ubiquitous transcription factors, including NFκB, Sp1 and c-Jun. The latter acts as a repressor of *TP53* expression: mouse fibroblasts from *cJUN*-deficient embryos have difficulties in progressing into the cell cycle due to suppression by an overexpressed p53 (Schreiber *et al.*, 1999). The p53 transcript is 2.8 kb long. The main initiation codon is in exon 2 but there is an alternative initiation site at codon 40, that leads to the production of a protein that lacks the N-terminus containing the transcription activation domain (DeltaNp53) (Courtois *et al.*, 2002). An alternatively spliced transcript that retains intron 2 has been identified (Ghosh *et al.*, 2004). This transcript is unable to encode the full-length protein but supports the synthesis of DeltaNp53. This feature is interesting since the same mode of regulation has been identified for *TP63* and *TP73*. Thus, production of DeltaN, transcription-deficient isoforms, seems to be a characteristic of *TP53* family members (Courtois *et al.*, 2004).

Post-translational regulation: signaling of DNA damage

The p53 protein is constitutively expressed in most cells and tissues as a latent factor. Due to its rapid turnover (5–20 minutes) the protein does not accumulate unless it is stabilized in response to a variety of intracellular and extracellular stimuli. Signals that activate p53 include diverse types of DNA damage (strand breaks, bulky adducts, oxidation of bases), blockade of RNA elongation, hypoxia, depletion of microtubules, ribonucleotides or growth factors, modulation of cell adhesion and alteration of polyamine metabolism (Pluquet and Hainaut, 2001). Most of our current knowledge of p53 protein activation is derived from studies using DNA strand breaks as inducing signals.

The main regulator of p53 protein activity is Mdm-2, a protein that binds p53 in the N-terminus (residues 17–29), conceals its transcription activation domain, redirects p53 from nucleus to cytoplasm and acts as an ubiquitin ligase to target p53 for degradation by the proteasome. The *MDM2* gene is a transcriptional target of p53, thus defining a feedback loop in which p53 controls its own stability (Moll and Petrenko, 2003). The p53/Mdm-2 complex is regulated by $p14^{Arf}$ (Alternative Reading Frame), a 14 kD protein encoded by an alternative reading frame of *CDKN2A*, the gene that encodes the tumor suppressor p16 (Moore *et al.*, 2003). The kinetics, extent and consequences of p53 activation vary according to the nature and intensity of the inducing signals. In response to ionizing radiation, activation of p53 proceeds through phosphorylation of p53 in the N-terminus by kinases involved in DNA-damage sensing, such as Atm (the product of the *Ataxia Telangiectasia* mutated gene) and Chk2 (a cell-cycle regulatory kinase) (Kastan *et al.*, 2001). Phosphorylation induces the dissociation of the p53/Mdm-2 complex, leading to p53 stabilization and allowing the binding of transcriptional co-activators, such as p300/CBP and pCAF, that acetylate p53 in the C-terminus. This process, as well as other, coordinated post-translational modifications of the C-terminus, induces conformational changes that turn the protein into an active form with high affinity for specific DNA response elements. The pathways of response to non-DNA damaging signals are less well understood and may involve different sets of phosphorylation events (Pluquet and Hainaut, 2001).

Downstream of p53: cell-cycle control, apoptosis and DNA repair

Once activated, p53 exerts its effects through two major mechanisms: transcriptional control (activation or repression of specific genes) and interference with the function of other proteins through

complex formation. Over 4800 genes have been identified as containing a p53-response element in regulatory regions (Wang *et al.*, 2001). Downstream effectors (Table 16.1) include regulators of cell-cycle checkpoints (in G1/S, G2 and during mitosis), factors involved in the signalling of apoptotic pathways (receptor- and/or mitochondria-dependent), genes mediating important aspects of DNA repair, and components of the transcription, replication and repair machineries. At the cellular level, activation of p53 generally induces either cell-cycle arrest (mostly in G1 and/or G2/M) or apoptosis. However, it must be realized that apoptosis is the preferred response in primary cells and that when cell-cycle arrest is induced, it is generally a permanent one, followed by cell senescence (Linke *et al.*, 1997). In other words, activation of p53 in a normal cell generally results in its permanent deletion from the pool of cells with proliferative capacity, providing a radical way for suppressing any cell that carries a risk of oncogenic transformation. How a cell ''selects'' between alternative responses (e.g. cell-cycle arrest or apoptosis) depends upon many factors, such as the nature and amplitude of the stress, as well as the cell type and differentiation status. In UV-irradiated skin, as well as in γ-ray irradiated intestine, the level and consequences of p53 activation are dependent upon cell differentiation status (Chaturvedi *et al.*, 2004; Wilson *et al.*, 1998). Together, these responses neutralize cells that carry the highest risk of propagating DNA-damage, while keeping alive other cells which are important in maintaining the integrity of the tissue. Thus, p53 can act as a mediator of not only genetic, but also tissue stability, in response to genotoxic stress.

Role of p53 in normal suppression of proliferation

The absence of major defects in *TP53*-deficient mice (aside from cancer susceptibility) has contributed to the misconception that the p53 protein may not exert any essential role in ''normal life'', its activity being restricted to protection from tumorigenic DNA damage. This, however, is an oversimplification. If studies in *TP53*-deficient mice clearly show that absence of p53 is not lethal, the presence of p53 may represent a very serious problem that cells need to overcome for normal development, cell growth and differentiation (Montes *et al.*, 1995). Spontaneous DNA damage occurs at rates of several thousand events per cell and per day, and p53 activation must be kept under tight control when the damage remains within the boundaries that can be safely handled by DNA repair systems. On the other hand, growth signaling itself activates p53 as a consequence of transcriptional activation of $p14^{Arf}$ by factors of the E2F family, which are at the receiving end of many signaling cascades stimulated by growth factors. Since $p14^{Arf}$ binds Mdm2 and releases it from interacting with p53, many proliferative signals have the potential to activate p53. This mechanism works as a safeguard against untimely and unprogrammed growth stimuli (Lomazzi *et al.*, 2002). Nevertheless, cells must keep it under control in order to allow normal cells to proceed into mitosis when requested.

Several mechanisms are likely to participate in the neutralization of p53 function when its activity is detrimental to normal cell life. For example, downregulation of p53 expression by c-Jun contributes to lower the levels of p53 protein at the G1/S transition, coinciding with the point where $p14^{Arf}$ is also expressed as a result of activation by E2F (Lomazzi *et al.*, 2002; Schreiber *et al.*, 1999). On the other hand, the DeltaNp53 isoform appears to accumulate during G1 and to reach a level that is higher than that of full-length p53 at the G1/S transition. Since DeltaNp53 lacks the transcription activation domain, it can work as a dominant negative inhibitor of p53 to allow cells to proceed into the cell-cycle (Courtois *et al.*, 2002). Overall, these mechanisms act in a co-ordinated manner as ''buffers'' against untimely p53 activation. Recently, two transcription regulators, YY1 and ELL, have been shown to negatively regulate p53 by preventing it from recruiting the transcription

Table 16.1. A selection of effectors of p53 functions

Factor	Activity	Mode of regulation	Function
Apo-1/Fas/CD95	Death signaling receptor	Transcriptional activation	Apoptosis
Bax-1	Dominant-negative inhibitor of bcl2	Transcriptional activation	Apoptosis
Bcl-2	Repressor of apoptosis through control of mitochondrial permeability	Transcriptional repression	Apoptosis
IGF-BP3	Inhibitor of IGF-I	Transcriptional activation	Apoptosis
Killer/DR5	Death signaling receptor	Transcriptional activation	Apoptosis
p85	Regulatory subunit of PI3 kinase	Transcriptional activation	Apoptosis
Pig-12	Glutathione transferase homologue	Transcriptional activation	Apoptosis
Pig-3	Quinone oxidase homologue	Transcriptional activation	Apoptosis
Pig-6	Proline oxidase homologue	Transcriptional activation	Apoptosis
IGF-I	Growth factor	Transcriptional repression	Apoptosis
IL-6	Survival factor	Transcriptional repression	Apoptosis
Thrombospondin-1	Inhibitor of angiogenesis	Transcriptional activation	Apoptosis
Noxa	Control of mitochondrial permeability	Transcriptional activation	Apoptosis
PUMA	Control of mitochondrial permeability	Transcriptional repression	Apoptosis
p3AIP	Control of mitochondrial permeability	Transcriptional repression	Apoptosis
Gadd45	Control of cell cycle in G2	Transcriptional activation	Cell cycle arrest, G2
BTG2	Inhibitor of proliferation	Transcriptional activation	Cell cycle arrest, G1
p21waf-1	Inhibitor of CDK2-4 and 6	Transcriptional activation	Cell cycle arrest, G1 and G2/M
Cyclin A	Cell-cycle regulation, S phase	Transcriptional repression	Cell cycle arrest, G1/S
13–3-3–σ	Control of Cdc25 at G2/M	Transriptional activation	Cell cycle arrest, G2/M
Cyclin G	Cell-cycle regulation	Transcriptional activation	Cell-cycle arrest?
GPx	Glutathione peroxidase	Transcriptional repression	Control of oxidative stress
NOS2/iNOS	Inducible nitric oxide synthase	Transcriptional repression	Control of oxidative stress
COX2	Inducible cyclooxygenase	Transcriptional activation	Anti-apoptotic effect?
Pig-1	Galectin-7	Transcriptional activation	Differentiation?
PCNA	Auxilliary subunit of polymerase δ	Transcriptional activation	DNA repair/replication
MSH2	Mismatch DNA Repair	Transcriptional activation	DNA repair
O6MGMT	O6-methylguaninemethyltransferase	Transcriptional activation	DNA repair
HOGG1	O8-deoxyguanosine glycosylase	Activation by protein binding	DNA repair
RPA	Replication protein A	Inhibition by protein binding	DNA repair/replication
ERCC2/ERCC3	Helicases, TFIIH complex	Activation by protein binding	DNA repair/ transcription
P53RR2	Ribonucleotide reductase homolog	Transcriptional activation	DNA repair?
TBP	TATA box-binding protein	Inhibition by protein binding	Inhibition of transcription
Mdm-2	Oncogene	Transcriptional activation	Repression of p53
MDR-1	Multidrug resistance	Transcriptional repression	Resistance to chemotherapy

co-activators p300/CBP. Both of these regulators are involved in differentiation processes. The YY1 (Ying-Yang 1) protein is a transcription factor that plays a role in bone morphogenesis and acts as a suppressor of p53 activation in response to genotoxic stress (Gronroos *et al.*, 2004). ELL (Eleven nineteen lysine rich leukemia) is a transcription elongation factor which can be rearranged by gene translocation to give rise to the MLL-ELL protein in acute lymphoblastic and myeloid leukemias. Both wild-type ELL and MLL-ELL can suppress normal p53 activity (Wiederschain *et al.*, 2003). This mechanism may explain why these cancers often develop in the presence of wild-type *TP53* alleles.

Why is *TP53* so frequently mutated in cancer?

Mutagenesis versus selection

TP53 is often inactivated by missense mutations, in contrast to many other tumor suppressors such as *APC*, *RB1*, *BRCA1* or p16/*CDKN2A* that are commonly inactivated by deletion or nonsense mutations (Olivier *et al.*, 2004). Of the 20 000 mutations reported to date, about 75% are single-base substitutions and most of them occur within exons encoding the DNA-binding domain (IARC *TP53* mutation database, www-p53.iarc.fr). Most of these mutations impair DNA-binding by disrupting the architecture of this domain or by eliminating contact points between the protein and target DNA (Cho *et al.*, 1994). About 30% of missense mutations affect six ''hotspot'' codons (175, 245, 248, 249, 273 and 282) (Figure 16.1). The other mutations are scattered over 200 different codons. Mutations occur through a number of biochemical pathways, including ''induced'' ones (that is, resulting from the direct attack of carcinogens at specific bases in the *TP53* sequence) and ''spontaneous'' ones (resulting from endogenous DNA modifications, imperfect DNA repair or polymerase errors during replication). Reading and interpreting mutation patterns in *TP53* to find out ''carcinogen fingerprints'' has become a favourite sport for many *TP53* mutation screeners (Hainaut and Hollstein, 2000; Olivier *et al.*, 2004).

There is no basis to the view that *TP53* lies in a ''hypermutable'' region of the genome. Thus, the high frequency and distribution of mutations reflects the acquisition of a selective advantage for tumorigenesis induced by the mutation. Indeed, inactivation of p53 function deletes a major cancer protection system by allowing cells with damaged DNA to replicate. Two main selection mechanisms can be proposed. The first accounts for the high rate of *TP53* mutations in tumors arising from intense exposure to environmental carcinogens such as, for example, lung cancer in smokers. In normal lung cells of heavy smokers, the high load of DNA damage over long periods of time generates a global suppression of cell growth, only compensated by the capacity of the cells to keep p53 function under control. When tobacco carcinogens hit within the *TP53* sequence to induce an inactivating mutation, the cell that has acquired such a mutation escapes stress-induced suppression and thus immediately gains a growth advantage against neighboring normal cells. This mechanism explains why the prevalence of *TP53* mutations in lung cancers of smokers shows a strict dose-dependence with tobacco consumption (Hainaut and Hallstein, 2000; Pfeifer *et al.*, 2002). In this model of carcinogenesis, *TP53* mutation is a very early, if not the earliest genetic modification that arises on the path to cancer. Indeed, high-sensitivity assays have shown that mutant *TP53* is detectable in non-cancer normal tissues of individuals exposed to exogenous or endogenous carcinogens (Hussain *et al.*, 2000; 2001).

The second selection mechanism acts through the $p14^{Arf}$-Mdm-2 connection. Through this connection, excessive activation of growth factor signalling (e.g through mutation of *RAS* genes, or constitutive activation of growth factor receptors) triggers p53 activation and growth suppression. In parallel with this mechanism, even limited hyperplasia in a solid tissue may lead to hypoxia and induction of p53 through yet another pathway

(Bardos and Ashcroft, 2004). Thus, only cells that contain a pre-existing *TP53* mutation will escape this control and will be allowed to proliferate under such critical conditions. In this context, a *TP53* mutation may appear, clinically, as a ''late'' event, and will be detectable only in lesions that have reached a certain stage in their progression. This is apparently the case for colorectal cancer developing according to a polyp–adenoma–carcinoma sequence, where *TP53* mutations become highly prevalent at the adenoma–carcinoma transition. It should be kept in mind, however, that the mutation may have occurred well ahead of this pathological transition, and becomes detectable in the lesion as a consequence of the expansion of a biologically selected clone. The selective advantage may result from enhanced survival in the poor physiological conditions which cancer cells are faced with during invasion (Guimaraes and Hainaut, 2002).

Beyond loss of suppression: gain-of-function effects

The paradigm of tumor suppressor genes implies that mutation results in inactivation, is accompanied by loss of the other allele, and the protein altogether disappears from cancer cells. With *TP53*, however, many missense mutations have a stabilizing effect and lead to the accumulation of mutant protein in the nucleus of cancer cells, sometimes in the presence of the remaining wild-type allele. This specificity has allowed the use of immunohistochemistry to detect ''mutant'' p53 in cancer tissues. While there is a fairly good correlation between missense mutation and protein accumulation, many mutations do not result in protein accumulation (e.g nonsense mutations, insertions and deletions) and wild-type p53 may accumulate in response to stress. Thus, care should be exercised in the use of immunohistochemistry as a surrogate for mutation analysis. The presence of high levels of mutant p53 and the frequent retention of a wild-type allele in cancer cells, has led to speculation that mutant p53 may exert dominant effects in carcinogenesis (de Vries *et al.*, 2002). Indeed, mutant p53 complexes with wild-type and inactivates its suppressor functions (dominant-negative effect). There is also some evidence that mutant p53 may behave as an independent, oncogenic protein. The molecular basis of this activity is not known, but the recent finding that mutant p53 can bind to and inactivate some p63 and p73 isoforms has led to the hypothesis that mutant p53 may also abrogate the suppressor functions of the other members of the family. This interaction might be dependent upon the polymorphism status at codon 72 of the *TP53* allele which contains the mutation (Bergamaschi *et al.*, 2003; Marin *et al.*, 2000).

Perspectives: p53 and cancer management

Over the past 15 years, many attempts have been made to translate basic knowledge of p53 protein structure and function into advances in cancer management and therapy (Table 16.2) (Seemann *et al.*, 2004). It is not over-pessimistic to state that many of these attempts have been unsuccessful in delivering clear and ready-to-use clinical solutions. This lack of success is due to many factors, but the most important is, in my view, the widespread tendency to underestimate the complexity of the biological processes in which p53 is involved. For example, all experimental evidence points to *TP53* gene status as a major determinant of cancer cell response to cytotoxic therapies (Fridman and Lowe, 2003). Since wild-type p53 is a key effector of apoptosis under conditions of DNA damage, it is widely expected that cancer cells that retain wild-type *TP53* allele are more sensitive to therapy than those containing mutant allele. This prediction holds true in many experimental systems. With real-life cancer, however, the *TP53* gene status is not a satisfying discriminator nor predictor of tumor response to treatment. This may be due to the fact that tumors with wild-type *TP53* alleles may have acquired other genetic alterations as an adaptation to the presence of wild-type p53. Another important confounding factor is that all

Table 16.2. Experimental therapy using p53 as a target

Agent (1)	Principle	Target (2)	Effect
PRIMA-1	Lipophilic compound, mechanism of action unknown	Cells with Mut p53	Reactivation of Mut p53, apoptosis
WR1065	Redox-activation of reactive thiols; enhances p53 binding to DNA	Wtp53	Chemo protective effects in normal cells
CDB3	Peptide binding to p53 core domain	Mut p53	Reactivation of Mut p53, apoptosis
CP-31398	''Adaptor'' binding to p53 core domain; stabilizes p53 into wt form	Mut p53	Apoptosis
ScFV421	Single chain antibody to C-terminal domain of p53	Wt and Mut p53	Activation of DNA-binding
ScFVME1	Single chain antibody to mut p53	Mut p53	Activation of transactivation
Adp53	Adenovirus expressing wtp53 for gene therapy	Cells with loss of p53 function	Restoration of wtp53 function in cancer cells
ONYX-15	E1B-deficient adenovirus that replicates only in p53-deficient cells	p53-deficient cells	Specific killing of cells that have lost p53 function

(1): For review see Seemann *et al.*, 2004.

(2): Mut p53: mutant p53 protein; wtp53: wild-type p53 protein.

mutant p53 proteins are not functionally equivalent. Under close examination, p53 mutants differ from each other by the degree of retention of wild-type properties, their dominant-negative effects, and their gain-of-function characteristics. More subtle classifications than ''yes'' or ''no'' for mutation are needed to fully exploit knowledge of *TP53* status for tumor prognosis (Martin *et al.*, 2002).

Detection of *TP53* mutation can be an aid in cancer diagnosis or prognosis, as recently shown in breast cancer (Olivier *et al.*, 2006). Several methods are now available, including in particular DNA chip-based methods, which resequence *TP53* with speed, sensitivity, cost and accuracy compatible with clinical use (Tonisson *et al.*, 2002; Wikman *et al.*, 2000). The detection of mutant *TP53* in a borderline, pre-neoplastic or cancer-predisposing lesion, should be seen as an indication that this lesion has high potential for evolution towards neoplasia, justifying more aggressive therapeutic approaches. Detection of mutant *TP53* can also be performed on histologically normal tumor margins as an additional test to evaluate the invasive capacity of a primary cancer. Mutations are sometimes detectable in surrogate specimens, such as DNA isolated from plasma or from exfoliated cells in subjects with suspicion of cancer but no identified, localized lesion. Finding of a mutation in such surrogate specimens should be taken as a justification for more intensive investigation in order to identify a primary lesion.

With respect to therapy itself, gene therapy aimed at reintroducing a wild-type *TP53* allele into cancer cells has shown only limited success (Seemann *et al.*, 2004). In addition to the difficulty of introducing the corrector gene into a sufficient number of target cells, it seems that cancer cells are extremely competent at eliminating the transgene. After all, they have already done it for their endogenous *TP53* alleles. Among other gene therapy approaches, one of the most impressive is the one developed by ONYX, based on the use of cytolytic viruses selectively replicating in *TP53*-deficient cells (McCormick, 2003). However, perhaps the most promising approach is the development of small lipophilic compounds or peptides that cross membranes and restore the

activity of mutant p53 proteins. The compound PRIMA-1, isolated by K. Wiman and collaborators in a screen of an NCI drug library, has shown remarkable potential to kill cancer cells which specifically express a mutant p53, but not their counterparts without p53 or expressing wild-type p53 (Bykov *et al.*, 2002*a*; Bykov *et al.*, 2002*b*). This compound has moderate toxicity in animal tests and may represent a model for developing drugs targeting p53 for human therapy. It will take years before such drugs are fully validated for clinical applications. It should be stressed, however, that there are pathological contexts in which simple drugs of low toxicity targeting mutant p53 may have a huge public health impact. In hepatocellular carcinomas that arise in subtropical areas of Asia and Africa, over half of the patients have a unique mutation in *TP53* at codon 249 as the result of mutagenesis by aflatoxins, a group of mycotoxins that are frequent contaminants of the diet in these areas (Szymanska *et al.*, 2004). There is currently no affordable therapy for these cancers in low-resource countries and the death toll of hepatocellular carcinomas is in the range of several hundreds of thousands per year. In such a bleak context, a small, simple drug against mutant p53 may make a world of difference and pave the way for extending this type of therapy to cancers where the *TP53* mutation pattern is more complex and heterogenous.

REFERENCES

Bardos, J. I. and Ashcroft, M. (2004). Hypoxia-inducible factor-1 and oncogenic signalling. *Bioessays*, **26**, 262–9.

Beckman, G., Birgander, R., Sjalander, A. *et al.* (1994). Is p53 polymorphism maintained by natural selection? *Hum Hered*, **44**, 266–70.

Bergamaschi, D., Gasco, M. and Hiller, L. (2003). p53 polymorphism influences response in cancer chemotherapy via modulation of p73-dependent apoptosis. *Cancer Cell*, **3**, 387–402.

Blandino, G. and Dobbelstein, M. (2004). p73 and p63: why do we still need them? *Cell Cycle*, **3**, 886–94.

Bykov, V. J., Issaeva, N., Selivanova, G. and Wiman, K. G. (2002*a*). Mutant p53-dependent growth suppression distinguishes PRIMA-1 from known anticancer drugs: a statistical analysis of information in the National Cancer Institute database. *Carcinogenesis*, **23**, 2011–18.

Bykov, V. J., Issaeva, N., Shilov, A. *et al.* (2002*b*). Restoration of the tumor suppressor function to mutant p53 by a low-molecular-weight compound. *Nat Med*, **8**, 282–8.

Chaturvedi, V., Qin, J. Z., Stennett, L., Choubey, D. and Nickoloff, B. J. (2004). Resistance to UV-induced apoptosis in human keratinocytes during accelerated senescence is associated with functional inactivation of p53. *J Cell Physiol*, **198**, 100–9.

Cho, Y., Gorina, S., Jeffrey, P. D. and Pavletich, N. P. (1994). Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations see comments. *Science*, **265**, 346–55.

Courtois, S., de Fromentel, C. C. and Hainaut, P. (2004). p53 protein variants: structural and functional similarities with p63 and p73 isoforms. *Oncogene*, **23**, 631–8.

Courtois, S., Verhaegh, G., North, S. *et al.* (2002). DeltaN-p53, a natural isoform of p53 lacking the first transactivation domain, counteracts growth suppression by wild-type p53. *Oncogene*, **21**, 6722–8.

de Vries, A., Flores, E. R., Miranda, B. *et al.* (2002). Targeted point mutations of p53 lead to dominant-negative inhibition of wild-type p53 function. *Proc Natl Acad Sci USA*, **99**, 2948–53.

Dumble, M. L., Donehower, L. A. and Lu, X. (2003). Generation and characterization of p53 mutant mice. *Methods Mol Biol*, **234**, 29–49.

Fridman, J. S. and Lowe, S. W. (2003). Control of apoptosis by p53. *Oncogene*, **22**, 9030–40.

Ghosh, A., Stewart, D. and Matlashewski, G. (2004). Regulation of human p53 activity and cell localization by alternative splicing. *Mol Cell Biol*, **24**, 7987–97.

Gronroos, E., Terentiev, A. A., Punga, T. and Ericsson, J. (2004). YY1 inhibits the activation of the p53 tumor suppressor in response to genotoxic stress. *Proc Natl Acad Sci USA*, **101**, 12165–70.

Guimaraes, D. P. and Hainaut, P. (2002). TP53: a key gene in human cancer. *Biochimie*, **84**, 83–93.

Hainaut, P. and Hollstein, M. (2000). p53 and human cancer: the first ten thousand mutations. *Adv Cancer Res*, **77**, 81–137.

Hanahan, D. and Weinberg, R.A. (2000). The hallmarks of cancer. *Cell*, **100**, 57–70.

Hussain, S.P., Amstad, P., Raja, K. *et al.* (2000). Increased p53 mutation load in noncancerous colon tissue from ulcerative colitis: a cancer-prone chronic inflammatory disease. *Cancer Res*, **60**, 3333–7.

Hussain, S.P., Amstad, P., Raja, K. *et al.* (2001). Mutability of p53 hotspot codons to benzo(a)pyrene diol epoxide (BPDE) and the frequency of p53 mutations in nontumorous human lung. *Cancer Res*, **61**, 6350–5.

Kastan, M.B., Lim, D.S., Kim, S.T. and Yang, D. (2001). ATM – a key determinant of multiple cellular responses to irradiation. *Acta Oncol*, **40**, 686–8.

Koster, M.I. and Roop, D.R. (2004). Transgenic mouse models provide new insights into the role of p63 in epidermal development. *Cell Cycle*, **3**, 411–13.

Linke, S.P., Clarkin, K.C. and Wahl, G.M. (1997). p53 mediates permanent arrest over multiple cell cycles in response to gamma-irradiation. *Cancer Res*, **57**, 1171–9.

Lomazzi, M., Moroni, M.C., Jensen, M.R., Frittoli, E. and Helin, K. (2002). Suppression of the p53- or pRB-mediated G1 checkpoint is required for E2F-induced S-phase entry. *Nat Genet*, **31**, 190–4.

Marin, M.C., Jost, C.A., Brooks, L.A. *et al.* (2000). A common polymorphism acts as an intragenic modifier of mutant p53 behaviour. *Nat Genet*, **25**, 47–54.

Martin, A.C., Facchiano, A.M., Cuff, A.L. *et al.* (2002). Integrating mutation data and structural analysis of the TP53 tumor-suppressor protein. *Hum Mutat*, **19**, 149–64.

Massion, P.P., Taflan, P.M., Jamshedur Rahman, S.M. *et al.* (2003). Significance of p63 amplification and overexpression in lung cancer development and prognosis. *Cancer Res*, **63**, 7113–21.

May, P. and May, E. (1999). Twenty years of p53 research: structural and functional aspects of the p53 protein. *Oncogene*, **18**, 7621–36.

McCormick, F. (2003). Cancer-specific viruses and the development of ONYX-015. *Cancer Biol Ther*, **2**, S157–S160.

McLure, K.G. and Lee, P.W. (1998). How p53 binds DNA as a tetramer. *EMBO J*, **17**, 3342–50.

Moll, U.M. and Petrenko, O. (2003). The MDM2–p53 interaction. *Mol Cancer Res*, **1**, 1001–8.

Montes, d.O.L., Wagner, D.S. and Lozano, G. (1995). Rescue of early embryonic lethality in Mdm2-deficient mice by deletion of p53. *Nature*, **378**, 203–6.

Moore, L., Venkatachalam, S., Vogel, H. *et al.* (2003). Cooperativity of p19ARF, Mdm2, and p53 in murine tumorigenesis. *Oncogene*, **22**, 7831–7.

Murray-Zmijewski, F., Lane, D.P and Bourdon, J.C. (2006). p53/p63/p73 isoforms: an orchestra of isoforms to harmonise cell differentiation and response to stress. *Cell Death Differ*, **13**, 962–72.

Olivier, M., Eeles, R., Hollstein, M. *et al.* (2002). The IARC TP53 database: new online mutation analysis and recommendations to users. *Hum Mutat*, **19**, 607–14.

Olivier, M., Hussain, S.P., Caron, d.F., Hainaut, P. and Harris, C.C. (2004). TP53 mutation spectra and load: a tool for generating hypotheses on the etiology of cancer. *IARC Sci Publ*, 247–70.

Olivier, M., Langerod, A., Carrieri, P. *et al.* (2006). The clinical value of somatic TP53 gene mutations in 1,794 patients with breast cancer. *Clin Cancer Res*, **15**, 1157–67.

Pfeifer, G.P., Denissenko, M.F., Olivier, M. *et al.* (2002). Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. *Oncogene*, **21**, 7435–51.

Pluquet, O. and Hainaut, P. (2001). Genotoxic and non-genotoxic pathways of p53 induction. *Cancer Lett*, **174**, 1–15.

Schreiber, M., Kolbus, A., Piu, F. *et al.* (1999). Control of cell cycle progression by c-Jun is p53 dependent. *Genes Dev*, **13**, 607–19.

Seemann, S., Maurici, D., Olivier, M., Caron de Fromentel, C. and Hainaut, P. (2004). The tumor suppressor gene TP53: implications for cancer management and therapy. *Crit Rev Clin Lab Sci*, **41**, 551–83.

Szymanska, K., Lesi, O.A., Kirk, G.D. *et al.* (2004). Ser-249TP53 mutation in tumour and plasma DNA of hepatocellular carcinoma patients from a high incidence area in the Gambia, West Africa. *Int J Cancer*, **110**, 374–9.

Taniere, P., Martel-Planche, G., Saurin, J.C. *et al.* (2001). TP53 mutations, amplification of P63 and expression of cell cycle proteins in squamous cell carcinoma of the oesophagus from a low incidence area in Western Europe. *Br J Cancer*, **85**, 721–6.

Tonisson, N., Zernant, J., Kurg, A. *et al.* (2002). Evaluating the arrayed primer extension resequencing assay of

TP53 tumor suppressor gene. *Proc Natl Acad Sci USA*, **99**, 5503–8.

Venot, C., Maratrat, M., Dureuil, C. *et al.* (1998). The requirement for the p53 proline-rich functional domain for mediation of apoptosis is correlated with specific PIG3 gene transactivation and with transcriptional repression. *EMBO J*, **17**, 4668–79.

Wang, L., Wu, Q., Qiu, P. *et al.* (2001). Analyses of p53 target genes in the human genome by bioinformatic and microarray approaches. *J Biol Chem*, **276**, 43604–10.

Wiederschain, D., Kawai, H., Gu, J., Shilatifard, A. and Yuan, Z. M. (2003). Molecular basis of p53 functional inactivation by the leukemic protein MLL-ELL. *Mol Cell Biol*, **23**, 4230–46.

Wikman, F. P., Lu, M. L., Thykjaer, T. *et al.* (2000). Evaluation of the performance of a p53 sequencing microarray chip using 140 previously sequenced bladder tumor samples. *Clin Chem*, **46**, 1555–61.

Wilson, J. W., Pritchard, D. M., Hickman, J. A. and Potten, C. S. (1998). Radiation-induced p53 and p21WAF-1/CIP1 expression in the murine intestinal epithelium: apoptosis and cell cycle arrest. *Am J Pathol*, **153**, 899–909.

Yang, A. and McKeon, F. (2000). P63 and P73: P53 mimics, menaces and more. *Nat Rev Mol Cell Biol*, **1**, 199–207.

17

Genetics of colorectal cancer

Susan M. Farrington and Malcolm G. Dunlop

Introduction

Colorectal cancer is a major public health problem in the developed world and is becoming increasingly prevalent in developing countries. The current annual world incidence is around 950 000 cases (Globocan, 2000). It is the most common cause of early cancer death in the non-smoking population. Recent developments have led to the isolation of a number of moderate- to high-risk cancer susceptibility genes for the disease. Identifying people with high-risk alleles offers real opportunities for application of preventive measures. Intensive surveillance to detect early cancer, or even prevent cancer by polyp removal, can be targeted by genotype information. Surgical intervention and chemoprevention guided by genetic information are also likely to be part of future armaments used to combat the disease. The last ten years has seen a number of exciting developments in understanding key molecular events involved in colorectal cancer susceptibility, which are beginning to provide new insight into the fundamental basis of the disease. In this chapter we will describe the major advances and how they are impacting diagnosis and clinical management of colorectal cancer.

Colorectal cancer epidemiology

The multifactorial etiology of colorectal cancer involves environmental factors as well as genetic susceptibility (see Chapter 14). There are large differences in global prevalence of the disease, which is generally four times higher in developed countries than in developing countries (IARC, WHO, 1997). Incidence rates also vary according to ethnicity (American Cancer Society, 2002), however the observed variation between countries is primarily due to the role of environmental factors. In support of this, there is a rising incidence of colorectal cancer in populations undergoing rapid economic development, in association with Westernisation of diet and lifestyle. Another example of environmental influence comes from migrant data; despite a historically relatively low colorectal cancer incidence in Japan, rates in Hawaiian Japanese are among the highest in the world (Parkin, 1992).

Figure 17.1a demonstrates that colorectal cancer is more common in males than in females and in both sexes the incidence rate increases with advancing age. This difference in sex incidence may partly be explained by evidence from a number of case-control and cohort studies which demonstrated an association between hormone replacement therapy (HRT) and colorectal cancer, with the majority of these suggesting a protective effect (Giacosa *et al.*, 1999). Accumulating evidence implicates obesity as a risk factor for colorectal cancer (Giacosa *et al.*, 1999), and a positive association may exist between colorectal cancer and diabetes (La Vecchia *et al.*, 1997). Not surprisingly, smoking has also been suggested as a significant risk factor (Giovannucci, 2001). Many studies have elucidated precise dietary and other variables responsible for the observed

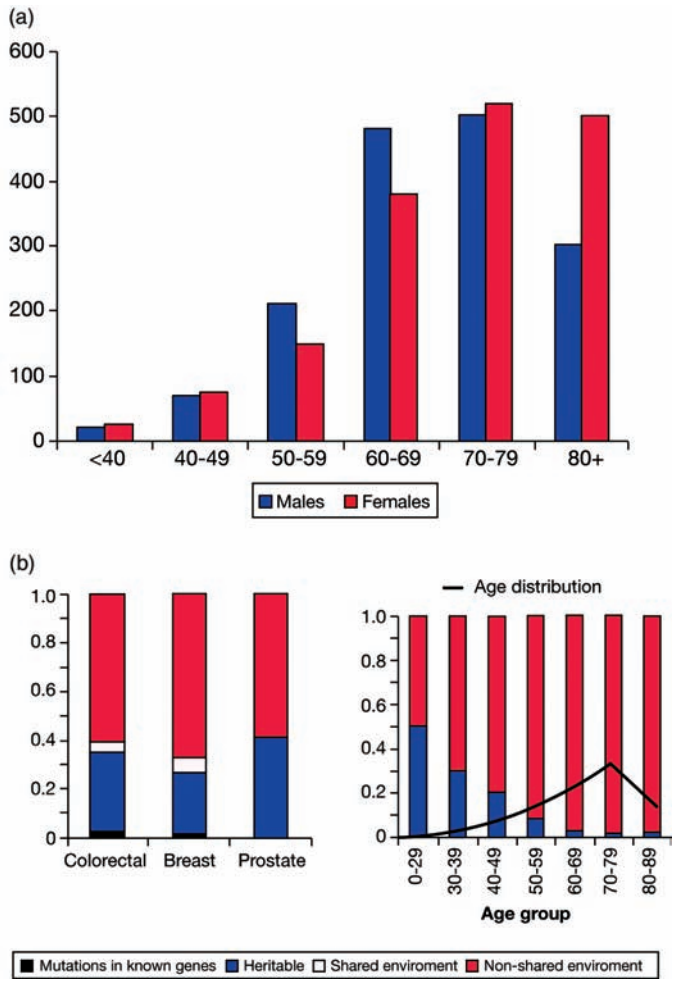

Figure 17.1 Colorectal cancer in the general population. a) Sex and age distribution of colorectal cancer. b) Heritable component of the disease: adapted information from Lichtenstein *et al.*, 2000 with incorporation of data for the known dominant gene mutations (left panel) and the age-dependent heritable component to colorectal cancer (right panel).

environmental influences on colorectal cancer incidence, suggesting that risk could be substantially reduced by diets high in vegetables and low in meat, together with regular physical activity and avoidance of alcohol (Potter, 1999; World Cancer Research Fund Panel, 2001). However, others have failed to provide evidence to uphold this hypothesis with clinical intervention studies (Alberts *et al.*, 2000; Schatzkin *et al.*, 2000), observational cohort studies (Fuchs *et al.*, 1999), as well as studies utilizing animal models (Rijnkels *et al.*, 1998; van Kranen *et al.*, 1998), showing no evidence of adenoma prevention related to diet. Nonetheless, adenoma formation is an intermediate phenotype and so results of further clinical studies with cancer prevention as the ultimate endpoint are awaited. One interesting variable, implicated as anti-tumor agents, is that of the family of non-steroidal anti-inflammatory drugs (NSAIDs). Both epidemiological evidence and experiments utilizing murine

models suggest that they may prevent colorectal cancer (Giardiello *et al.*, 1993; Steinbach *et al.*, 2000) and many studies are now investigating the cellular mechanisms by which these drugs act, with the aim of developing novel chemopreventive agents (Stark and Dunlop, 2005).

Genetic susceptibility to colorectal cancer

Although environmental factors are clearly important in the etiology of the disease, there is a significant input from genetic factors. First-degree relatives of colorectal cancer cases are considered to have a 2–4-fold increased risk, with around 20–25% of all colorectal cancer cases being associated with a family history (Bonelli *et al.*, 1988). Early screening studies suggested that cancer susceptibility is due to a predisposition in the development of colorectal adenomas (Cannon-Albright *et al.*, 1988). The population frequency of the ''adenoma-prone'' allele was calculated as 19% but not all adenomas progress to cancer and dietary effects might clearly influence the expression of such an allele in terms of both adenoma and cancer. A North European twin study has now provided good evidence that about 35% of all colorectal cancer cases have a genetic component (Lichtenstein *et al.*, 2000) (Figure 17.1b), which is intriguing as known susceptibility alleles account for only 5% of all cases. The genetic factors involved are poorly understood and it is now clear that colorectal cancer susceptibility is complex. Recent evidence has shown that recessive inheritance of susceptibility alleles can also play a major role in increased colorectal cancer risk (Farrington *et al.*, 2005), and the hunt is on for new risk alleles of moderate to low penetrance, or associated with a marginal increased risk, and for those involved in complex gene–gene and gene–environment interactions.

Of the known genetic factors, there are two autosomal dominant and one autosomal recessive inherited cancer syndromes that account for a significant minority of all colorectal cancer cases (~5%). Familial Adenomatous Polyposis (FAP) is a rare syndrome caused by mutations in the Adenomatous Polyposis Coli (*APC*) gene and characterized by the presence of multiple adenomas. In other multiple adenoma families, bi-allelic mutations of the Base Excision Repair (BER) gene *MUTYH* were identified as the causative lesions and the gene has been associated with increased colorectal cancer risk. The major colorectal cancer syndrome is hereditary non-polyposis colorectal cancer (HNPCC). Affected kindreds have a high occurrence of colorectal and certain extra-colonic cancers and early age of onset. HNPCC has no pathognomonic features and has traditionally been diagnosed on the basis of family history. However, due to the lack of definitive biomarkers for HNPCC, these diagnostic criteria have been pragmatic and are not inclusive of all cases. Clinical criteria such as the Amsterdam criteria will only identify families where the gene defect is highly penetrant and the families are of sufficient size to allow the appropriate number of cases to arise. Thus there is a bias that tends to exclude small families inappropriately. The underlying defect in the majority of HNPCC has been identified as lesions in the DNA mismatch repair (MMR) pathway.

Familial adenomatous polyposis coli (FAP)

FAP is a highly penetrant autosomal dominant disorder, clinically characterised by the development of hundreds to thousands of adenomatous polyps of the colon and rectum. This rare syndrome has an annual incidence of around 1:7000 live births (Bisgaard *et al.*, 1994). However malignancy is virtually inevitable if surgery is not undertaken and thus accounts for ~0.2% of all colorectal neoplasia. There are a number of extra-colonic features such as multiple craniofacial and long bone osteomata, epidermoid cysts, retinal pigmentation, gastroduodenal polyposis and malignancy, desmoid tumors and an increased risk of peri-ampullary, papillary thyroid, brain tumors and sarcomas.

FAP germline mutations

The gene responsible for FAP was identified via cytogenetic localization and genetic linkage studies to chromosome 5q21–22 (Bodmer *et al.*, 1987; Leppert *et al.*, 1987). The Adenomatous Polyposis Coli (*APC*) gene was cloned and germline mutations identified in a number of FAP patients (Kinzler *et al.*, 1991; Joslyn *et al.*, 1991). *APC* comprises an 8.5 kb transcript encoding a 2843 amino acid polypeptide in 15 exons. The majority of germline mutations reported in FAP patients/families result in premature truncation of APC (Nagase and Nakamura, 1993; Mandl *et al.*, 1994). Around 80% of the *APC* mutations identified to date are in exon 15, with 2 specific mutations (codons 1061 and 1309) accounting for 15–20% of all *APC* mutations but the remainder are spread throughout the gene with no other ''hotspots'' (Cottrell *et al.*, 1992; Groden *et al.*, 1993). Short repeat sequences at the amino terminus of APC are predicted to form coiled-coil structures suggesting that normal APC functions as a homodimer. Thus mutations leading to a truncated APC protein may result in a heterodimer of mutant/wild-type APC protein which may abrogate the function of the normal protein in a dominant-negative manner (Su *et al.*, 1993).

Phenotype–genotype correlations can be implied from the type or location of the *APC* gene mutation and these may help inform the clinical treatment for the patient and whether extracolonic manifestations will occur, such as desmoid disease. Indeed the common codon 1309 mutation is associated with a dense polyp phenotype and a high cancer risk in the retained rectum, thus favoring certain surgical interventions over others (Nagase *et al.*, 1992). However, these correlations are not well defined with various groups identifying families with identical *APC* mutations but diverse phenotypes in terms of both colorectal polyposis and extracolonic disease (Paul *et al.*, 1993). A number of variant syndromes are also caused by germline *APC* mutations. Gardner's syndrome encompasses florid polyposis with epidermoid cysts of skin, osteomas of mandible and congenital hypertrophy of the retinal pigment epithelium. Attenuated FAP (AAPC or AFAP) describes patients with later onset disease and greatly reduced polyp numbers of a slightly flatter type. The *APC* mutations responsible for this phenotype tend to be found upstream of codon 200 and downstream of codon 1600. An association with desmoid disease has been correlated to 3′ *APC* gene mutations (Eccles *et al.*, 1996). Turcot's syndrome manifests as multiple colonic adenomas with early onset of colorectal cancer and also tumors of the central nervous system, particularly brain tumors. If the associated brain tumor is cerebellar medulloblastomas the genetic defect is likely to be in the *APC* gene, whereas if glioblastoma multiforme tumors are found, the defect is likely to be defective DNA mismatch repair (Hamilton *et al.*, 1995). Although general phenotype–genotype correlations are being drawn, there are still many variables to be considered, especially in view of potential modifier loci.

Cellular functions of the APC protein

The function of the *APC* gene product is a focus of intense research and the complexities of the cellular role of APC have yet to be fully understood. The APC protein is expressed in epithelial cells in the upper portions of the colonic crypts, suggesting involvement in colonocyte maturation (Smith *et al.*, 1993; Nathke *et al.*, 1996). Several functional domains are revealed in the protein sequence including the N terminal homodimerization sequences. Other domains indicate that APC is involved in numerous cellular processes including cellular adhesion, cell-cycle regulation, apoptosis, differentiation and intracellular signal transduction. The central portion of the APC protein contains β-catenin binding and regulation domains and binding domains for the axin family of proteins. APC may influence cellular adhesion by affecting the interaction between catenins and E-cadherin, thus promoting the shedding and migration of epithelial cells. In conjunction with other proteins such as axin, glycogen synthase

kinase 3β (GSK) and other GSK binding proteins, APC plays a critical role in intracellular communication by modulating the levels of β-catenin dependent transcription (cellular role of APC reviewed in Barth *et al.*, 1997 and Ben-Ze'ev and Greiger, 1998). β-catenin is an important transcription factor for oncogenic proteins such as cyclin D1 and c-myc (He *et al.*, 1998). Extensive studies are underway in model organisms to fully define the APC/β-catenin signaling pathway, (Figure 17.2a), but the pivotal role of this pathway in colorectal tumorigenesis is exemplified by the identification of somatic mutations in many of the components, such as APC, β-catenin and axin, with somatic *APC* mutations being identified in ~85% of all colorectal tumors (Polakis, 1999). Thus all these genes are candidate colorectal cancer susceptibility alleles, however only germline mutations of the *APC* gene have been commonly identified. A germline mutation of the axin homolog *Axin2* was identified as being involved in tooth agenesis, however there was also a correlative predisposition to colorectal neoplasms (Lammi *et al.*, 2004) but other studies have not supported a germline role for *Axin2* in colorectal disease (Peterlongo *et al.*, 2005).

Other functions of wild-type APC involve the microtubule cytoskeleton. In vitro studies indicate that wild-type APC binds to microtubules and promotes their formation and bundling (Munemitsu *et al.*, 1994; Nathke *et al.*, 1996). The organization and structure of microtubules are vital to cell division and migration and truncated forms of the APC protein appear to be unable to bind microtubules (Smith *et al.*, 1994). Not only does APC bind to microtubules but it presents them to the kinetochores (Kaplan *et al.*, 2001; Fodde *et al.*, 2001), the protein complexes that mediate attachment of chromosomes to the spindle apparatus in order to accurately separate the sister chromatids during mitosis. Mouse cells homozygous for a truncating mutation, Apc^{min}, displayed abnormal chromosome patterns when compared to wild-type counterparts. This evidence indicates that the APC protein plays an important role in maintaining fidelity of chromosome segregation and thereby contributes to the prevention of aneuploidy. This underscores an important observation that aneuploidy occurs in the majority of colorectal cancers and indicates further complexity in the role that *APC* mutations play in tumorigenesis and progression. The carboxy terminus of the APC protein not only binds microtubules but also the microtubule binding protein EB1 (Su *et al.*, 1995) and the tumor suppressor protein Dlg (Matsumine *et al.*, 1996), which are both implicated in tumorigenesis. Once the underlying function of APC has been clearly determined it may be possible to restore or augment the effects of deleterious mutations by pharmacological means.

Modifiers of FAP

Overwhelming evidence exists for modifier genes affecting the phenotypic expression of this disease, from the wide spectrum of clinical presentation of FAP, even within families where affected individuals all carry the same *APC* mutation (Burt, 1996). Mouse models for FAP have been generated and the Apc^{min} mouse (Moser *et al.*, 1990) phenotype is modulated by an unlinked locus (Moser *et al.*, 1992) which was identified as the gene *Mom1* (Dietrich *et al.*, 1993) encoding a secretory phospholipase A2 (MacPhee *et al.*, 1995), a protein involved in lipid metabolism and prostaglandin production. The homologous human gene was found to be located on chromosome 1p35–36 where evidence suggests the presence of a modifier locus (Tomlinson *et al.*, 1996), however no correlation for phospholipase A2 gene mutation with observed phenotype has been established and thus it is unlikely to be a modifier of FAP (Dobbie *et al.*, 1996). Another candidate mouse modifier loci has been identified in *Apc* mutant mice, which may encode genes imparting X-ray responsiveness (van der Houven van Oordt *et al.*, 1999). The *Apc* mouse models allow specific investigation of gene–environment interactions and also novel treatments for preventing the development of colorectal neoplasia in FAP. Dietary and other environmental effects are also being investigated.

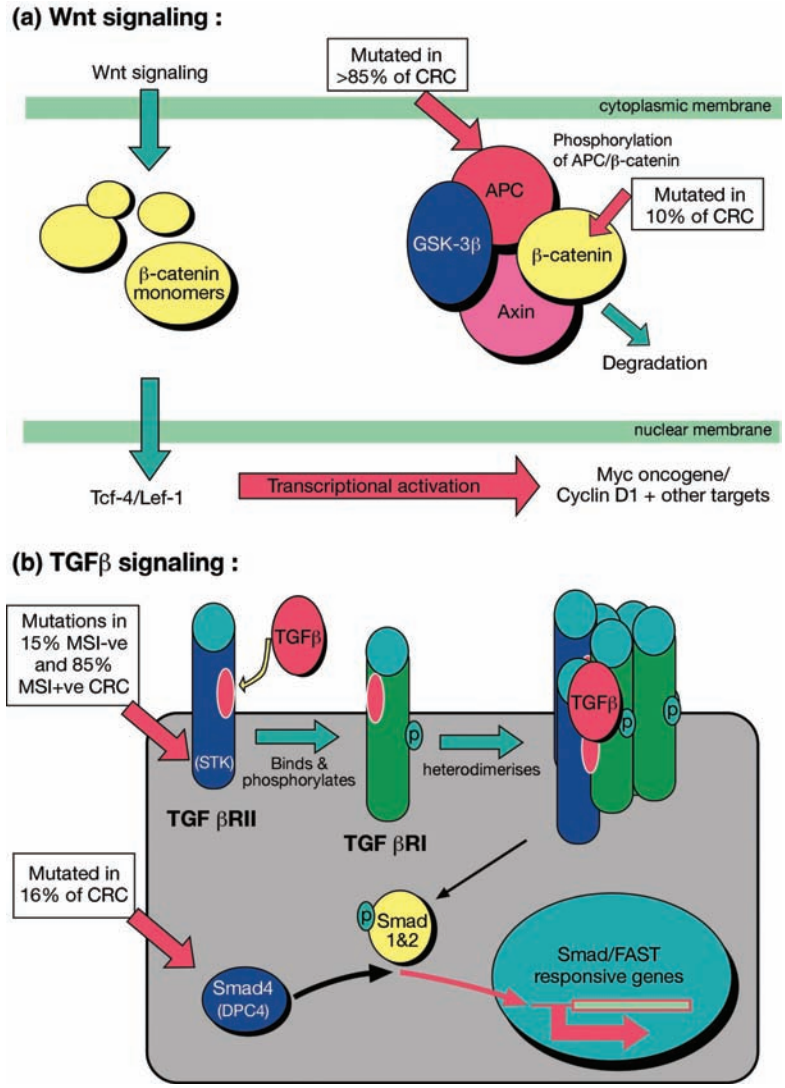

Figure 17.2 Signalling pathways commonly dysregulated in colorectal cancer. a) The Wnt signalling cascade controlling β-catenin levels, is dysregulated either by mutation in the *APC* gene (~85% of cases) and in a large proportion of the other cases, β-catenin itself is mutated. This ultimately leads to increases of β-catenin levels and its subsequent translocation into the nucleus and transcriptional activation of genes such as c-myc and cyclinD1. b) The TGF β signalling pathway is involved in growth regulation and may of its components are implicated in colorectal carcinogenesis as both germline and somatic events.

One study demonstrated that *Apc* mutant mice fed a Western-style diet (high fat, low calcium and vitamin D), had an increase in tumor incidence (Yang *et al.*, 1998).

Future research into understanding the structure, function and interactions of the APC protein, combined with the use of mouse models, will undoubtedly lead to new treatments of not only

FAP but also sporadic colorectal cancer. Identification of the *APC* gene was the start of an exciting avenue of research that will have considerable relevance to those affected with colorectal cancer.

Hereditary non-polyposis colorectal cancer (HNPCC) or lynch syndrome

HNPCC, or the eponymous Lynch syndrome, is an autosomal dominant disorder with high penetrance in which colorectal cancer develops in gene carriers but without the numerous adenomas seen in FAP (Lynch *et al.*, 1985). Classically, HNPCC was defined as the development of carcinomas in the proximal part of the colon, however, expression of the HNPCC phenotype is far more diverse in terms of age of onset and the organs and site where malignancy can develop: besides colorectal cancer, the trait can be associated with uterine, gastric, ovarian, upper urinary tract, small intestinal and other malignancies. The recognition that there is an increased risk of other cancers, in fact a greater risk for endometrial cancer in females than colorectal cancer, means that the term Lynch syndrome has gained popularity.

Strict criteria such as the Amsterdam criteria have been used to clinically define HNPCC and this has a major effect on the apparent prevalence of HNPCC and also considerably influences the apparent penetrance of the gene defect(s). However, current systematic assessment of the genes responsible, on a whole population basis, will allow the true prevalence and penetrance to be determined.

There are no robust biomarkers for HNPCC. There is a predominance of certain tumor histological features in HNPCC, such as a lymphoid response, mucinous histology and poor differentiation (Jass *et al.*, 1994), but these features are not specific and do not accurately identify all cases of familial colorectal cancer. The most useful biomarker for HNPCC has been identified as a result of elucidation of the molecular basis of this syndrome, namely tumor genetic instability or MSI (microsatellite instability) due to defective DNA mismatch repair (MMR). While not all MSI+ tumors are due to germline defects in MMR genes, patients with tumors exhibiting this phenotype are highly enriched for HNPCC. Systematic analysis of MMR genes and the prevalence of MSI+ tumors, is allowing the overall contribution of these genes to cancer susceptibility to be determined.

DNA mismatch repair genes

Identification of colorectal tumors displaying widespread genomic instability at short repetitive DNA tracts suggested that defective MMR might be involved, in view of previous work in bacteria and yeast (Grilley *et al.*, 1990; Strand *et al.*, 1993; Ionov *et al.*, 1993; Aaltonen *et al.*, 1993; Thibodeau *et al.*, 1993). Whilst observations of a mutator phenotype in HNPCC patients' tumors were being made, a systematic linkage mapping strategy identified a locus linked to the anonymous marker, D2S123, on the short arm of chromosome 2, in two families (Peltomaki *et al.*, 1993). A second locus was identified by a report of linkage to a chromosome 3p marker shortly afterwards (Lindblom *et al.*, 1993), clearly establishing locus heterogeneity in HNPCC.

Positional cloning and candidate gene isolation strategies identified the human homologues of the yeast and bacterial MMR genes. The first was the MutS homolog *MSH2* on chromosome 2p (Fishel *et al.*, 1993; Leach *et al.*, 1993). Truncating and splicing mutations were identified and shown to co-segregate with the disease in HNPCC families. As yeast strains containing the MutL homologs MLH or PMS were also found to demonstrate repetitive tract instability (Strand *et al.*, 1993), the human homologs of these genes were obvious candidates and mutations were demonstrated in the HNPCC families (Bronner *et al.*, 1994; Papadopolous *et al.*, 1994). *MLH1* was shown to be the gene segregating with chromosome 3p markers and *PMS1* and *PMS2* lie on chromosomes 2q and 7q respectively (Nicolaides *et al.*, 1994); *MSH6* (or GTBP (Jiricny *et al.*, 1988)) has been implicated in colorectal cancer susceptibility (Miyaki *et al.*, 1997) however *MSH3* (Watanabe

et al., 1996) is only thought to be involved somatically in tumor formation. A fourth MutL homolog which induces microsatellite instability when mutated has been identified, *MLH3* (Lipkin *et al.*, 2000), however it has more homology to the bacterial and yeast proteins, than to mammalian MMR proteins, and its role in HNPCC is questionable (Hienonen *et al.*, 2003), due to the functional redundancy likely between PMS2/PMS1 and MLH3. It is possible that, like *MSH6*, it may play a role in the less typical HNPCC cancer families.

MMR mutation spectrum and genotype–phenotype correlations

The locus heterogeneity of HNPCC is compounded by mutation heterogeneity in the MMR genes. Germline mutations identified include base substitutions, and insertions/deletions both short in length and large genomic changes including whole exons. The proportion of HNPCC families due to mutations in each gene remains to be elucidated but data so far suggests that around 53% of deleterious mutations identified are in *MLH1*, 42% in *MSH2* and 4% in *MSH6* [InSiGHTat www.insight-group.org]. The other MutL homologs *PMS2*, *PMS1* and *MLH3* have very few reported mutations, probably due to redundancy of the proteins in the pathway. Germline changes in *PMS2* have been implicated as part of a recessive syndrome as described below, predisposing to unusual malignancy spectrums such as those found in Turcot's syndrome (de Rosa *et al.*, 2000; de Vos *et al.*, 2004). Furthermore there are some consanguineous families where offspring have bi-allelic inactivation for MLH1 and their phenotype is much more severe with the development of early onset hematological maligancies (also prevalent in mouse models deficient for MMR) and Neurofibromatosis Type 1 (Ricciardone *et al.*, 1999; Wang *et al.*, 1999). MSI analysis of somatic tissues from the children demonstrated that there was indeed a constitutional mismatch repair deficiency in the normal tissues. One other report describes a compound heterozygote with *MLH1* missense mutations where the patient developed breast cancer at the age of 35 (Hackman *et al.*, 1997).

Genotype–phenotype correlations are being formulated with reports that *MSH2* mutations are associated with an excess risk of extra-colonic cancers such as transitional cell carcinoma of the renal-pelvis and ureter, as well as carcinoma of the stomach and ovaries (Vasen *et al.*, 1996). *MSH2* and *MLH1* mutations impart a substantially elevated risk of endometrial and small bowel cancer (Vasen *et al.*, 1996; Dunlop *et al.*, 1997). Other reports suggest that gastric cancer is more prevalent in *MLH1* than in *MSH2* gene carriers (11% vs. 4.5%) and that these cancers are a specific histological type associated with environmental etiology (Aarnio *et al.*, 1995). *MSH6* gene mutations are associated with atypical HNPCC families that have an excess of endometrial cancers, a delayed age of onset of the disease and incomplete penetrance (Akiyama *et al.*, 1997; Wijnen *et al.*, 1999). Our own data indicates that gender also influences cancer risk: the lifetime risk of colorectal cancer in *MSH2*/*MLH1* gene carriers was significantly greater for males than females 74% vs. 30%. The endometrial risk in females was 42%, and the lifetime risk for all cancers was 91% for males and 69% for females.

Another rare autosomal dominant cancer susceptibility syndrome called Muir–Torre syndrome has been found to be an allelic variant of HNPCC and is diagnosed by the presence of at least one sebaceous gland neoplasm (adenoma, epithelioma or carcinoma) and/or a keratoacanthoma and at least one internal malignancy. Germline mutations in either *MSH2* or *MLH1* have been identified in Muir–Torre syndrome but other genetic and environmental factors may play a role in the differences in phenotypic expression (Kolodner *et al.*, 1994; Bapat *et al.*, 1996). Germline mutations of *PMS2* or *MLH1* are also associated with Turcot's syndrome as described above.

There is no clear link between phenotype and specific germline mutations at any given gene despite considerable allelic heterogeneity. Missense mutations in the MMR genes may result in a less severe phenotype or lower penetrance, as functionally relevant structural changes may be

less severe (Beck *et al.*, 1997). Indeed, more recent data supports this, with an attenuated allele of *MLH1* increasing sporadic colorectal cancer risk (Lipkin *et al.*, 2003). Further studies of the relationship of mutation to clinical outcome using in vitro systems (Shimodaira *et al.*, 1998; Trojan *et al.*, 2002) and also animal models (Yang *et al.*, 2004) may help in this regard.

Mismatch repair defects in humans and DNA instability in tumors

One useful marker identified in HNPCC is tumor instability of repetitive tract DNA in microsatellite markers (Ionov *et al.*, 1993; Aaltonen *et al.*, 1993) and in tumor-promoting genes such as *TGFβRII*, *BAX* and *APC* (Markowitz *et al.*, 1995; Huang *et al.*, 1996; Ouyang *et al.*, 1998). Approximately 16% of sporadic colorectal cancers show MSI and tend to be right-sided, diploid, show an inverse relationship with p53 mutation and are associated with good prognosis (Lothe *et al.*, 1993). The major mechanism of somatic inactivation of DNA mismatch repair in sporadic cases, is epigenetic silencing of MLH1 by hypermethylation of CpG islands in the promoter (termed CpG island methylator phenotype) (Cunningham *et al.*, 1998; Toyota *et al.*, 1999). Thus deficiency of the DNA mismatch repair process can occur via different mechanisms; germline mutation and subsequent inactivation of the wild-type allele, or, in sporadic cases, by two somatic alterations, hypermethylation of *MLH1* being the most common.

Identification of microsatellite instability in tumors on a prospective basis may have utility in understanding the influence of these different mechanisms on clinicopathological factors. Studies of early-onset colorectal cancer, demonstrate that more than half display the MSI phenotype (Liu *et al.*, 1995) and that the majority of these contain identifiable germline MMR gene mutations (Farrington *et al.*, 1998). However we have also demonstrated that inference of data from older cohorts to early-onset cases may not be entirely valid (Farrington *et al.*, 2002): tumor histopathology was significantly different between the two cohorts and MSI was only associated with better prognosis in the older cohort. This indicates that the different molecular mechanisms of the MSI phenotype may play a significant role in survival. Other studies investigating MSI in colorectal cancer patients diagnosed at under the age of 50 (Gryfe *et al.*, 2000) suggest that MSI is an independent predictor of a favorable outcome in this age group. However, no germline analysis of MMR genes was performed and so this may represent fundamental age-dependent differences in tumor MSI origin. Further studies are required to clarify the impact of germline mutation on prognosis.

Microsatellite instability can be detected in a number of different cancers from HNPCC gene carriers (Aaltonen *et al.*, 1994) and occurs in tumors from patients with multiple different primary cancers (Horii *et al.*, 1994), suggesting that such individuals may well carry DNA repair gene mutations even without a family history of cancer. Another complexity of MMR inactivation was a report of a germline epimutation in two patients with multiple cancers, presenting evidence that hypermethylation of *MLH1* could actually be inherited (Suter *et al.*, 2004).

Molecular mechanisms of colorectal carcinogenesis and defective MMR

The human MMR system appears to be similar to yeast with *MSH*, *MLH* and *PMS* homologs. A homolog of the E. coli *MutH* gene has not been detected in yeast but has been recently reported for the human system. Bellacosa *et al.* (1999) demonstrated that DNA repair protein MED1/MBD4 (a methyl CpG binding protein with homology to bacterial DNA repair glycosylases/lyases) forms a complex with MLH1 and displays endonuclease activity. It has been suggested that MED1 is the human equivalent of the bacterial endonuclease, MutH. Mismatch repair involves recognition and binding of mismatches by two separate MSH heterodimers. MutSα is a complex of MSH2/MSH6 and binds G:T mismatches and small deletions/insertions, whereas MutSβ is a complex of MSH2/MSH3 and binds looped DNA denoting

larger deletions and insertions (Fishel, 2001). The correct strand for excision is recognized and the MutL complexes work to direct excision of a 1–2 kb segment of DNA and subsequent repair by DNA polymerase. Further mismatch repair systems are likely to be identified. A nucleotide-specific mismatch repair system that recognizes deaminated bases such as 5MeC which produce G:T mismatches (Wiebauer *et al.*, 1990) would be of great interest, as they result in C–T transitions which are frequently seen in the *APC* gene in colorectal tumors. Intriguingly, recent evidence has implicated one of the DNA (cytosine-5)-methyltransferases, *Dnmt1*, in MMR (Guo *et al.*, 2004). The other predominant *APC* gene tumor mutations are short deletions or insertions at repetitive sequences and an interesting hypothesis is that many of the somatic changes involved in the genesis of colorectal cancer may be caused by defects in DNA MMR. Indeed, in tumors displaying MSI, we have observed a marked propensity for somatic mutations of the *APC* gene to occur at repeat tracts (Huang *et al.*, 1996). However, other studies have not observed any particular predominance of such mutations and this may be due to differences in the proportion of tumors arising in germline MMR gene mutation carriers. Nonetheless, it is clear that MMR deficiency results in an elevated mutation rate in several key genes involved in colorectal carcinogenesis (reviewed in Grady, 2004).

MMR gene mouse models have been created but inactivation generally results in the development of lymphomas rather than gastrointestinal tumors (reviewed in Wei *et al.*, 2002). Interestingly, the PMS1 mouse model does not display any tumor burden and a number of the other MMR knockouts have meiotic phenotypes resulting in reduced fertility or infertility, highlighting the role of these genes in gametogenesis. These models are helping our understanding of the other functions of MMR genes, which include methylation tolerance and hyper-recombination between divergent sequences (de Wind *et al.*, 1995), a role in the cell-cycle checkpoint mechanisms (Koi *et al.*, 1994) and also in apoptosis (Zhang *et al.*, 1999). MMR-deficient cells showing resistance to alkylating agents (methylation tolerance), are often resistant to other cytotoxic drugs such as 5-fluoro-uracil (Carethers *et al.*, 1999) in a p53 dependant manner (Bunz *et al.*, 1999). These drugs are commonly used in the treatment of colorectal cancer patients but may have genotoxic effects on normal tissues of patients with a MMR defect. Thus, future treatment regimes may be tailored according to the genetic profile of a given tumor with respect to MMR status.

Much work remains to be done in understanding the effects of these HNPCC gene mutations in carcinogenesis, their effect on mutation rate in the homozygous and heterozygous state, on population gene frequency, penetrance and the question of other DNA repair systems that may be involved in heritable predisposition to colorectal cancer.

Other candidate susceptibility genes involved in HNPCC

Not all HNPCC families are due to mutations of known MMR genes, suggesting that defects in other, genes may be important. There is intense research interest in fully defining the deranged cellular pathways involved in colorectal cancer at the somatic level and many of the components of these pathways are potential candidate susceptibility alleles.

The TGFβ/SMAD pathway is an interesting example (Figure 17.2b) as its disruption plays a role in sporadic and familial colorectal cancer. Germline *TGFβRII* mutations occur in a small fraction of atypical colorectal cancer families and were found to segregate with disease, resulting in disruption of TGFβ signalling (Lu *et al.*, 1998). Recently a variant in the *TGFβRI* gene has been identified as a high-frequency, low-penetrant risk allele for colorectal cancer (Kaklamani *et al.*, 2003). Unravelling the complexities of *TGFβ* signaling has resulted in characterization of the molecular basis of another colorectal cancer predisposition syndrome, namely Juvenile Polyposis (see below).

Chromosome 18q was first suggested to harbor a tumor suppressor gene involved in colorectal cancer by frequent cytogenetic deletions and loss of heterozygosity in that region. Subsequently a number of genes were cloned, including *DCC* (deleted in colorectal cancer (Fearon *et al.*, 1990)), *JV18* and *DPC4* (Deleted in Pancreatic Cancer (or *SMAD*4) (Hahn *et al.*, 1996)). Heterozygous mutations of the *SMAD4* gene in mice induce a cancer phenotype, although it is a fairly late age of onset (Sirard *et al.*, 1998). Germline mutations of *SMAD4/DPC4* have been identified in a number of Juvenile Polyposis familes (Howe *et al.*, 1998) and will be discussed in the next section. This authenticates the idea that disruption of the *TGFβ* signaling pathway has important implications for colorectal carcinogenesis and emphasizes the complex interrelationship of genes involved in colorectal neoplasia at the somatic and germline level.

Rare dominant disorders

In addition to FAP and HNPCC, there are a number of rarer syndromes that are associated with a smaller elevated risk of colorectal cancer. Each of the syndromes are characterized by benign polyposis of the intestine and the polyps are of unusual histology. The clinical phenotypes of these syndromes are heterogeneous, and there is some clinical overlap between them.

Juvenile Polyposis syndrome (JPS)

Juvenile Polyposis syndrome is an autosomal dominant syndrome, with incomplete penetrance and manifests as diffuse hamartomatous polyps of the colon, small bowel and stomach, with onset at a very early age (under 10 years). Genetic linkage studies led to analysis of candidate genes on chromosome 18q and germline mutations of *SMAD4/DPC4* were demonstrated. A small number of JPS families have also been associated with mutation in the *PTEN* gene, the gene found to be responsible for Cowden disease (CD; see below) which may reflect the clinical overlap between these two disorders (Olshwang *et al.*, 1998). The complexities of the relationship of cancer susceptibility and the *TGFβ/SMAD* signaling pathway have been outlined above. A minority of JPS cases are also due to mutations in BMPR1A, another component of the SMAD/FAST pathway (Howe *et al.*, 2001). This emphasizes again that understanding the pathway involved can lead to identification of other susceptibility loci that might have been difficult to uncover using exclusively genetic approaches.

Peutz-Jeghers syndrome (PJS)

Peutz-Jeghers syndrome is an autosomal dominant condition with low penetrance. It is characterized by the development of hamartomatous polyposis of the gastrointestinal tract and in 95% of the patients the hallmark feature is melanin spots on the lips and buccal mucosa. There is an increased risk of gastrointestinal, pancreas, testis, ovaries, breast and uterus cancers (Giardello *et al.*, 1987). The PJS gene was mapped to chromosome 19p13.3, cloned and mutations identified in a number of families (Hemminki *et al.*, 1998). The gene encodes a novel serine/threonine kinase and was named *LKB1/STK11*. *LKB1* is involved in signaling and regulation of cellular differentiation (Jenne *et al.*, 1998). Given the paradigms of the APC/β-catenin and the *TGFβ/SMAD* pathways, it seems likely that other components of this pathway may be involved in colorectal tumorigenesis and perhaps susceptibility. However, to date no other genes in the pathway have been implicated in tumorigenesis (Alhopuro *et al.*, 2005).

Cowden disease (CD)

Cowden disease is a rare autosomal dominant disorder characterized by the development of oral and facial papules in concert with hamartomatous lesions of the thyroid, breast, skin and digestive tract and occasionally with neurological features. Mapping to chromosome 10q22–24 (Nelen *et al.*, 1996) led to the identification of *PTEN* in which

mutations were identified in the families (Liaw *et al.*, 1997). *PTEN* encodes a putative tyrosine phosphatase that is downregulated by *TGFβ*. *PTEN* also shares homology to the focal adhesion protein tensin and is therefore implicated in many different regulatory roles such as cell motility, migration, communication etc., which on disruption lead to an increase in invasion and metastasis (Li and Sun, 1997). For both JPS and CD, the genes identified do not account for all the families and thus more genes are likely to be identified as susceptibility genes for these syndromes. Indeed, genetic mapping in a Hereditary Mixed Polyposis family has suggested a locus on chromosome 6q, which as the phenotype is similar to JPS may play a role in both syndromes (Thomas *et al.*, 1996).

Recessive inheritance

The last few years have provided concrete evidence for the first recessive syndromes predisposing to colorectal neoplasia. The first was suggested in Turcot's syndrome families, where the underlying defect was found to be bi-allelic mutations of *PMS2*, a DNA mismatch repair gene (see HNPCC section) (de Rosa *et al.*, 2000), however it is unclear as to whether bi-allelic inheritance is always required, although the true situation is complicated by the presence of several pseudogenes of *PMS2* (de Vos *et al.*, 2004). Secondly, investigation of multiple adenoma families who were negative for *APC* germline mutations, yielded a new mechanism by which colorectal neoplasia may progress. Analysis of adenoma and tumor material for somatic *APC* mutations identified a specific mutation spectrum consisting of G:T transversions, a hallmark of defective BER. The BER gene *MUTYH* was identified as the recessive susceptibility gene involved in multiple adenoma formation (Al-Tassan *et al.*, 2002).

MYH-associated colorectal neoplasia

A number of studies in multiple adenoma families and FAP families negative for *APC* mutations have been found to contain homozygous or compound heterozygous variants of the *MUTYH* gene (Sieber *et al.*, 2003; Sampson *et al.*, 2003; Venesio *et al.*, 2004). The protein product, MYH, is an integral part of cellular DNA repair systems with a major role in the repair of 8oxoG:A mispairs. Interestingly, a major source of these mispairs is through oxidative damage, which is prevalent in the gut. Hence determining the relevance of *MUTYH* as a colorectal cancer risk allele is of great importance. A number of studies have shown that colorectal cancer cases display bi-allelic disruption of the *MUTYH* gene (Enholm *et al.*, 2003; Wang *et al.*, 2004), however some recent studies have been able to show true association of *MUTYH* variants with colorectal disease (Croitoru *et al.*, 2004; Farrington *et al.*, 2005). Case-control studies have allowed the excess risk imparted by a specific genotype and the penetrance of the disease to be estimated. Bi-allelic inactivation of MYH predisposes to a highly penetrant phenotype, with almost complete penetrance by age 60 and an increased genotype relative risk (RR) of developing the disease by >90 fold (Farrington *et al.*, 2005).

MYH protein function and interactions

Studies into the function of the BER proteins have been fuelled by its newly identified role in colorectal neoplasia (Cheadle and Sampson, 2003). The role of BER in removing oxidative damage is complex and there are several proteins employed to remove the damaged nucleotide. Studies of MYH have implicated it in the role of long patch BER (Parker *et al.*, 2001) and in replication coupled repair (Boldogh *et al.*, 2001). However, MYH function is generally implied from its *E.coli* homolog, *MutY*, which recognizes and removes the adenine from post-replicative mispairs of 8-oxo-G:A, G:A or C:A (David and Williams, 1998). Besides the nucleus, MYH also localizes to mitochondria, probably due to the prevalence of oxidation occurring there (Ohtsubo *et al.*, 2000). One interesting area is the interaction with other repair pathways, including DNA MMR, which is also defective in a proportion of colorectal cancer

(see HNPCC above). The MutSα complex is not only activated by oxidative damage (Mazurek *et al.*, 2002) but has been shown to interact with MYH via the MSH6 protein (Gu *et al.*, 2002).

Common genetic variation and susceptibility to colorectal cancer

The genes responsible for syndromic colorectal cancer susceptibility make up only a minority of cases. It is highly likely that there remain a number of rare high penetrance alleles yet to be discovered, as evidenced by the existence of large families segregating high penetrance alleles such as one locus identified on chromosome 15q (CRAC) where a common risk haplotype was identified (Jaeger *et al.*, 2003). It will only be a matter of time before the causative mutation is identified and other smaller families and cancer series can be tested for other alleles of the gene responsible for the linkage signal at the CRAC locus.

Notwithstanding these high penetrance alleles, the majority of the unexplained genetic contribution to colorectal cancer is likely to be due to a number of low-risk alleles (Kemp *et al.*, 2004). However, as will become clear from the discussion below, only a very few alleles are reasonably well supported by statistically robust experimentation and replication studies. The evidence in most cases is inconclusive and much remains to be done to fully evaluate the role of these variants where there is suggestive evidence for association. It is fair to say that there are no variants for which the current weight of evidence is overwhelming. The problems to date have been lack of sufficient size of case-control studies, subgroup analysis, publication bias, multiple testing (or ''data-mining''), lack of replication between different population groups, admixture and population stratification. However, there are a number of variants that are worthy of comment with respect to colorectal cancer risk. Although alleles conferring RR of around 1.5 may not seem useful in the context of clinical genetics, such effects can provide novel insight into disease causation and lead to study of new pathways found to be aberrant in cancer. Furthermore, a RR of 1.5 has considerable public health relevance because screening for colorectal cancer in the whole population is currently being recommended. Thus, targeted screening and more intensive approaches to those at higher risk could potentially radically improve the efficiency of such public health measures.

Non-truncating variants in the *APC* gene

As discussed above, the vast majority of truncating variants in the *APC* gene cause a polyposis phenotype and are associated with a very high colorectal cancer risk. However, in view of its central gatekeeper role, *APC* remains a strong candidate contributing to colorectal cancer susceptibility through common genetic variation. A number of non-truncating alleles have been described and two *APC* variants have sufficient weight of evidence to merit consideration as possible risk modifying alleles; these being the E1317Q variant and the D1822V variant in association with dietary fat intake.

Another interesting observation arising through the study of FAP families provides the best evidence that common genetic variation contributes to colorectal neoplasia, that of the ''pre-mutation'', I1307K. I1307K is a germline variant that results in an unstable poly-A_8 tract within the *APC* gene (Laken *et al.*, 1997). The variant causes genomic instability in the *APC* gene and a marked increased frequency of somatic frameshift *APC* mutations in colonic epithelial cells. The resultant somatic inactivating *APC* mutations result in an attenuated polyposis phenotype and an excess colorectal cancer risk. However, although the variant is common in Ashkenazi Jews (6%), it is rare in other populations. Nonetheless, such pre-mutations may be more common than previously considered in cancer genetics and will considerably complicate the detection of risk alleles, since they in themselves do not appear to be pathogenic.

There is conflicting evidence that the germ-line missense variant, E1317Q, confers a risk for colonic adenomatous polyps (Lamlum *et al.*, 2000).

The population prevalence of the allele is around 2–3% and some evidence suggests it is associated with multiple colorectal adenomas with an odds ratio (OR) of 11 and accounting for around 4% of all patients with multiple colorectal adenomas. However, contradictory data from subsequent studies mean that these findings remain controversial. One recent study found no excess of E1317Q heterozygotes among patients with colorectal cancer or adenomatous polyps compared to the general population. Although, when allele frequencies were compared between disease groups and colonoscopically normal subjects, there appeared to be a significant excess of the risk allele in the colorectal cancer patients (Hahnloser *et al.*, 2003). However, the issue of subgroup analysis and lack of stringent matching by age/sex and environmental risk factor exposure remains a problem. Hence, the contribution of E1317Q to large bowel neoplasia requires investigation in very large case series with stringently matched controls to definitively address the issue.

Another variant in APC, D1822V, may be associated with a protective effect from large bowel cancer. There is tentative evidence that homozygotes for valine at the 1822 residue have a reduced risk of colon cancer (OR, 0.6; 95% CI, 0.4–1.0) for those aged over 65 years. However, there was a marked and highly significant effect when combining a low-fat diet and the protective allele (OR, 0.2; 95% CI, 0.1–0.5) relative to those who were homozygous wild-type and consumed high-fat diets. These results provide strong evidence that individuals with the valine/valine variant are at reduced risk of colon cancer if they eat a low-fat diet (Slattery *et al.*, 2001). However, again it is important that replication studies are undertaken as many studies involving a first observation involve considerable data mining and so the issue of multiple testing comes to the fore.

Cyclin D1 (CCND1)

Aberrant expression of the cell cycle gene, cyclin D1 (*CCND1*), is a common somatic event in colorectal cancer but there is conflicting evidence as to whether genetic variation contributes to colorectal cancer susceptibility. There is some data to suggest that *CCND1* variants may modify the phenotype of HNPCC (Kong *et al.*, 2001) but there have been no substantial replication studies to confirm or refute the findings of this small study. Similarly, there remains controversy as to whether *CCND1* variants may be risk alleles of major effect themselves. Reported putative associations should be taken with circumspection because the apparent association has only been detectable in certain patient subgroups (Kong *et al.*, 2001; Le Marchand *et al.*, 2003). Furthermore, other reports suggest that there is no effect (Grieu *et al.*, 2003) and so the jury remains out as far as the involvement of *CCND*1 in colorectal cancer susceptibility is concerned.

Methyl tetrahydrofolate reductase (MTHFR)

Methylenetetrahydrofolate reductase is centrally involved in regulation of folate metabolism, and thus indirectly in DNA synthesis and repair. Epidemiological evidence has suggested that dietary folate consumption is inversely related to colorectal cancer risk. Common variants in the *MTHFR* gene, C677T and A1298C variants may be associated with an increased cancer risk in association with low dietary folate. The C677T variant influences enzyme activity, while the A1298C variant may be in linkage disequilibrium (LD) with C677T, or both may be in LD with a separate pathogenic variant (Kemp *et al.*, 2004). Thus, there is considerable rationale for studies investigating variation in *MTHFR*.

However, despite the substantial background rationale, the evidence to date remains inconclusive that genetic variants in *MTHFR* are involved in colorectal cancer susceptibility (Sharp and Little, 2004). At best, the evidence is only suggestive of a moderately reduced colorectal cancer risk associated with the 677TT genotype, especially when folate intake is high. However there appears to be a complex relationship between alcohol

consumption, folate and genotype. Once again, several studies have undertaken subgroup analysis and appear to demonstrate an effect for some risk groups, but it is a clear that there is a need for large population-based association studies in which risk exposure and dietary history is systematically recorded in order to determine whether or not genetic variation in *MTHFR* plays a role in colorectal cancer susceptibility.

STK15 *(Aurora-A)*

STK15 is involved in mitotic chromosomal segregation and there is considerable evidence suggesting aberrant gene function is involved in neoplastic transformation and progression. A genetic variant resulting in an amino acid substitution, F31I, was identified through mapping a cancer susceptibility locus in outbred mice (Ewart-Toland *et al.*, 2003). The variant has been shown to have cellular transformation capacity and to be associated with tumor aneuploidy in colorectal cancer. In a meta-analysis of case-control studies, there is robust evidence for a significantly increased risk of colorectal cancer in homozygotes, with an OR of 1.5 (Ewart-Toland *et al.*, 2005). This is supported by biological rationale since aneuploidy and chromosomal instability is a hallmark of the majority of colorectal cancers. Interestingly the gene resides in a chromosomal area showing allelic imbalance in Finnish familial colorectal cancer cases (Laiho *et al.*, 2003). Thus, understanding the fundamental basis of cancer, such as aneuploidy in colon cancer cells, can allow convergence of data from different research strategies and provide evidence to support the causal role of particular genes in cancer susceptibility, as well as providing reassurance that data generated in animal models and in vitro systems have relevance to human cancer. Indeed, *STK15* is only one of many genes involved in mitotic chromosomal segregation, and it is likely that findings likely only provide the first of many mechanisms leading to understanding of the role of tumor cell aneuploidy in association with cancer susceptibility.

Cytochrome P-450

Cytochrome P-450 is involved in activation and detoxifying many carcinogenic substances, especially those present in smoke. Furthermore, cigarette smoking is related to increased risk of colorectal adenomas and of rectal cancer, especially in men. Polycyclic aromatic hydrocarbons are activated by cytochrome P4501A1 (CYP1A1) and detoxified by glutathione S-transferases. There are a number of genetic polymorphisms in genes involved in the metabolism of polycyclic aromatic hydrocarbons and other tobacco-related carcinogens and so these are reasonable candidate susceptibility loci, especially when interacting with environmental factors. It has thus been suggested that variation in cytochrome genes might be involved in cancer susceptibility. The evidence for involvement remains inconclusive with several studies showing an apparent effect and others none. The largest and most recent study of almost 2000 cases of colorectal cancer and over 2000 controls (Slattery *et al.*, 2004) suggests that the *CYP1A1* genotype alone is not associated with colon or rectal cancer. However, in combination with a risk *GSTM1* allele and an allele associated with a *NAT2* rapid-acetylator phenotype, there was a significantly increased risk of colon cancer (OR = 1.7), especially amongst men who smoke (OR = 2.5). A similar result was noted for male smokers with respect to rectal cancer. This large study provides substantial evidence for a contribution of *GSTM1*, *CYP1A1* and *NAT2* alleles, although replication studies are needed. Nonetheless, the data support the association between smoking and large bowel malignancy as well as indicating a genetic modifying effect through *CYP1A1* and *GSTM1* genotypes.

*Cell-cycle checkpoint kinase 2 (*CHEK2*)*

CHEK2 is involved in cellular responses to DNA damage that culminate in cell cycle arrest, apoptosis, and/or DNA repair. The 1100delC variant of *CHEK2* has been shown to be causally involved in a

small proportion of breast cancer families and acts as a low penetrance susceptibility allele for breast cancer, with a twofold increased breast cancer risk for carriers (Meijers-Heijboer *et al.*, 2002). In some breast cancer families there is an excess of colorectal cancer and so by inference, *CHEK2* has been proposed as a colorectal cancer susceptibility gene. However, although descriptive studies in breast and colorectal cancer families seem to indicate a common genetic etiology, namely the *CHEK2* 1100delC allele, the largest association study to date has failed to show any convincing effect (Kilpivaara *et al.*, 2003). A recent study has suggested an effect in certain risk subgroups (de Jong *et al.*, 2005) but again large studies are clearly required to determine definitively whether *CHEK2* variants confer any excess risk of colorectal cancer, the level of any associated risk, and the overall contribution of such alleles to colorectal cancer disease burden. In short, the role of *CHEK2* in colorectal cancer susceptibility remains unresolved.

Other putative risk alleles

In addition to those discussed above, there are a substantial number of putative colorectal cancer risk alleles that have been proposed with varying degrees of support (Kemp *et al.*, 2004). We discuss a few of these in outline below in order to provide a flavor of those which appear to have at least some support for impacting on colorectal cancer risk.

A common length polymorphism in the gene encoding the type 1 receptor for TGFβ, TGFβR1*6A, has been extensively studied in a number of cancers and a meta-analysis performed (Kaklamani *et al.*, 2003). The evidence for a role in colorectal cancer susceptibility is inconclusive since the effect appears to be related to population with an OR of 1.38 in the data from studies of populations living in the USA, while there was no effect for European populations. Nonetheless, components of the TGF signaling pathway constitute strong candidate genes in view of the antiproliferative and apoptosis-promoting role of TGFβ. Furthermore, if confirmed in larger population-based studies, the TGFβR1*6A polymorphism could have considerable public health importance as the carrier frequency is ~15%, implying an overall contribution of 1.2% to the total burden of colorectal cancer, greater than that of FAP and approaching that of HNPCC alleles.

Loss of imprinting of the insulin-like growth factor II gene (*IGF2*) has been found in the normal colonic mucosa of around a third of patients with colorectal cancer, a threefold excess in comparison with controls (Cui *et al.*, 2003). Patients with colorectal cancer had a significantly greater loss of *IGF2* imprinting (OR = 21.7) in circulating lymphocytes, indicating it may be a marker for colorectal cancer. However, further evidence also indicates that not only is loss of *IGF2* imprinting heritable but it is also associated with increased tumor burden in an animal model (Sakatani *et al.*, 2005). Thus it is possible that epigenetic mechanisms are also involved in colorectal cancer susceptibility, further complicating attempts to identify the genes contributing to the disease burden.

Apolipoprotein E (apoE) is involved in bile acid, cholesterol and triglyceride metabolism. Variants in apoE have been associated with a number of diseases and recently their involvement in cancer has been proposed. Since bile acids have been causally implicated in colorectal cancer etiology, it is certainly plausible that colorectal cancer risk might be modified by apoE genotype. One sizeable study suggest that the E3 apoE allele was associated with a protective effect, with around 30% excess risk for those not carrying the protective allele, especially those aged 65 years or older (Slattery *et al.*, 2005).

There are many other putative risk alleles that have been identified in various populations, including insulin-like growth factor-1 (IGF-1), its binding protein (IGFBP), polymorphisms in *COX2* genes and in other genes encoding components of the arachidonic acid metabolism pathway, a length variant in the *H-ras* gene, variants in GSTT1, and a C/T variant 13910 bp upstream of the lactase coding sequence in the Finnish population. However, each of these candidates, however

plausible the gene or pathway, requires to be tested in large case-control series and then replicated in other populations, again with large case and control series. The size of the primary studies should be numbered in thousands given likely effects and the problems of multiple testing with similar numbers used to undertake the validation and replication studies. Only then can a reasonable view be taken on the likely involvement of common variation in large bowel malignancy.

Conclusions

Identifying genes responsible for autosomal dominant colorectal cancer susceptibility syndromes such as HNPCC and FAP has resulted in major advances in patient management, translating directly into a reduction in death rate from malignancy for gene carriers. Thus these syndromes serve as a paradigm for high-penetrance colorectal cancer susceptibility alleles. Clearly, it is incorrect to surmise that such dramatic benefits to the individual will result from identifying lower-penetrance alleles. However, major public health benefits will undoubtedly result from such discoveries and it is likely that worthwhile benefits for the individual may also result. However, more importantly, discovering even low-penetrance alleles will afford novel insights into disease causation, opening up a new discipline of rational drug development aimed at chemoprevention and risk reduction, much as is the case for modern prevention of coronary heart disease.

As will be apparent from the discussion above, the problem at present is that each of the associations that have been reported are not well supported by robust statistics and study design, or with replication studies in similarly large collections of intensively phenotyped cases and controls. As new technologies come online, that allow whole-genome approaches to be utilized toward understanding the complex genetics of colorectal cancer, many more putative candidates will be identified. Hence, it is important that resource and effort is not wasted by undertaking underpowered association studies. What is needed are very large population-based studies in which the epidemiological design is both robust and comprehensive. We await the results of such studies with considerable anticipation.

Acknowledgements

The work undertaken by SMF and MGD and the Colon Cancer Genetics Group is supported by grants from Cancer Research UK (C348/A3758), Medical Research Council (G0000657–53203), Scottish Executive (K/OPR/2/2/D333, CZB/4/41, CZB/4/94), Association for International Cancer Research (02–330).

REFERENCES

Aaltonen, L.A., Peltomaki, P., Leach, F.S. *et al.* (1993). Clues to the pathogenesis of familial colorectal cancer. *Science*, **260**, 812–16.

Aaltonen, L.A., Peltomaki, P., Mecklin, J.-P. *et al.* (1994). Replication errors in benign and malignant tumours from hereditary non-polyposis colorectal cancer patients. *Cancer Research*, **54**, 1645–8.

Aarnio, M., Mecklin, J.-P., Aaltonen, L.A. *et al.* (1995). Life-time risk of different cancers in hereditary nonpolyposis colorectal cancer (HNPCC) syndrome. *International Journal of Cancer*, **64**, 430–3.

Akiyama, Y., Sato, H., Yamada, T. *et al.* (1997). Germ-line mutations of the *hMSH6/GTBP* gene in an atypical hereditary nonpolyposis colorectal cancer kindred. *Cancer Research*, **57**, 3920–3.

Alberts, D.S., Martinez, M.E., Roe, D.J. *et al.* (2000). Lack of effect of a high-fiber cereal supplement on the recurrence of colorectal adenomas. Phoenix Colon Cancer Prevention Physicians' Network. *New England Journal Medicine*, **342**, 1156–62.

Alhopuro, P., Katajisto, P., Lehtonen, R. *et al.* (2005). Mutation analysis of three genes encoding novel LKB1-interacting proteins, BRG1, STRADα, and MO25α, in Peutz-Jeghers syndrome. *British Journal of Cancer*, **92**, 1126–9.

Al-Tassan, N., Chmiel, N.H., Maynard, J. *et al.* (2002). Inherited variants of MYH associated with somatic

G:C–T:A mutations in colorectal tumors. *Nature Genetics*, **30**, 227–32.

American Cancer Society. (www.cancer.org). *Cancer facts and figures 2002*.

Bapat, M., Xia, L., Madlensky, L. *et al.* (1996). The genetic basis of Muir–Torre syndrome includes the hMLH1 locus. *American Journal of Human Genetics*, **59**, 736–9.

Barth, A.I.M., Nathke, I.S. and Nelson, W.J. (1997). Cadherins, catenins and APC protein: interplay between cytoskeletal complexes and signaling pathways. *Current Opinion in Cell Biology*, **9**, 683–90.

Beck, N.E., Tomlinson, I.P.M., Homfray, T. *et al.* (1997). Genetic testing is important in families with a history suggestive of hereditary non-polyposis colorectal even if the Amsterdam criteria are not fulfilled. *British Journal of Surgery*, **84**, 233–7.

Bellacosa, A., Cicchillitti, L., Schepis, F. *et al.* (1999). MED1, a novel human methyl-CpG-binding endonuclease, interacts with DNA mismatch repair protein MLH1. *Proceedings of the National Academy of Sciences of the USA*, **96**, 3969–74.

Ben-Ze'ev, A. and Geiger, B. (1998). Differential molecular interactions of β-catenin and plakoglobin in adhesion, signaling and cancer. *Current Opinion in Cell Biology*, **10**, 629–39.

Bisgaard, M.L., Fenger, K., Bulow, S. *et al.* (1994). Familial adenomatous polyposis (FAP): frequency, penetrance, and mutation rate. *Human Mutation*, **3**, 121–5.

Bodmer, W.F., Bailey, C.J., Bodmer, J. *et al.* (1987). Localisation of the gene for familial adenomatous polyposis on chromosome 5. *Nature*, **328**, 614–16.

Boldogh, I., Milligan, D., Lee, M.S. *et al.* (2001). hMYH cell cycle-dependent expression, subcellular localization and association with replication foci: evidence suggesting replication-coupled repair of adenine:8-oxoguanine mispairs. *Nucleic Acids Research*, **29**, 2802–9.

Bonelli, L., Martines, H., Conio, M. *et al.* (1988). Family history of colorectal cancer as a risk factor for benign and malignant tumours of the large bowel. A case control study. *International Journal of Cancer*, **41**, 513–17.

Bronner, C.E., Baker, S.M., Morrison, P.T. *et al.* (1994). Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. *Nature*, **368**, 258–61.

Bunz, F., Hwang, P.M., Torrance, C. *et al.* (1999). Disruption of p53 in human cancer cells alters the responses to therapeutic agents. *Journal of Clinical Investigation*, **104**, 263–9.

Burt, R.W. (1996). Familial risk and colorectal cancer. *Gastroenterology Clinics of North America*, **25**, 793–803.

Cahill, D.P., Lengauer, C., Yu, J. *et al.* (1998). Mutations of mitotic checkpoint genes in human cancers. *Nature*, **392**, 300–3.

Cannon-Albright, L.A., Solnick, M.H., Bishop, D.T. *et al.* (1988). Common inheritance of susceptibility to colonic adenomatous polyps and associated colorectal cancers. *New England Journal of Medicine*, **319**, 533–7.

Carethers, J.M., Chauhan, D.P., Fink, D. *et al.* (1999). Mismatch repair proficiency and in vitro response to 5-fluorouracil. *Gastroenterology*, **117**, 123–31.

Cheadle, J.P. and Sampson, J.R. (2003). Exposing the MYtH about base excision repair and human inherited disease. *Human Molecular Genetics*, **12**, R159–R165.

Cottrell, S., Bicknell, D., Kaklamanis, L. *et al.* (1992). Molecular analysis of *APC* mutations in familial adenomatous polyposis and sporadic colon carcinomas. *Lancet*, **340**, 626–30.

Croitoru, M.E., Cleary, S.P., DiNicola, N. *et al.* (2004). Association between biallelic and monoallelic germline *MYH* gene mutations and colorectal cancer risk. *Journal of the National Cancer Institute*, **96**, 1631–4.

Cui, H., Cruz-Correa, M., Giardiello, F.M. *et al.* (2003). Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. *Science*, **299**,1753–5.

Cunningham, J.M., Christensen, E.R., Tester, D.J. *et al.* (1998). Hypermethylation of the h*MLH1* promoter in colon cancer with microsatellite instability. *Cancer Research*, **58**, 3455–60.

David, S.S. and Williams, S.D. (1998). Chemistry of glycosylases and endonucleases involved in base-excision Repair. *Chemical Reviews*, **98**, 1221–61.

de Jong, M.M., Nolte, I.M., Te Meerman, G.J. *et al.* (2005). Colorectal cancer and the *CHEK2* 1100delC mutation. *Genes Chromosomes Cancer*, **43,** 377–82.

De Rosa, M., Fasano, C., Panariello, L. *et al.* (2000). Evidence for a recessive inheritance of Turcot's syndrome caused by compound heterozygous mutations within the *PMS2* gene. *Oncogene*, **19**, 1719–23.

De Vos, M., Hayward, B.E., Picton, S. *et al.* (2004). Novel *PMS2* pseudogenes can conceal recessive mutations causing a distinctive childhood cancer syndrome. *American Journal of Human Genetics*, **74**, 954–64.

de Wind, N., Dekker, M., Berns, A. *et al.* (1995). Inactivation of the mouse *Msh2* gene results in mismatch repair deficiency, methylation tolerance, hyperrecombination, and predisposition to cancer. *Cell*, **82**, 321–30.

Dietrich, W.F., Lander, E.S., Smith, J.S. *et al.* (1993). Genetic identification of Mom-1, a major modifier locus affecting Min-induced intestinal neoplasia in the mouse. *Cell*, **75**, 631–9.

Dobbie, Z., Muller, H. and Scott, R. J. (1996). Secretory phospholipase A2 does not appear to be associated with phenotypic variation in familial adenomatous polyposis. *Human Genetics*, **98**, 386–90.

Dunlop, M. G., Farrington, S. M., Carothers, A. D. *et al.* (1997). Cancer risk associated with germline DNA mismatch repair gene mutations. *Human Molecular Genetics*, **6**, 105–10.

Eccles, D. M., van der Luijt, R., Breukel, C. *et al.* (1996). Hereditary desmoid disease due to a frameshift mutation at codon 1924 of the *APC* gene. *American Journal of Human Genetics*, **59**, 1193–201.

Enholm, S., Hienonen, T., Suomalainen, A. *et al.* (2003). Proportion and phenotype of *MYH*-associated colorectal neoplasia in a population-based series of Finnish colorectal cancer patients. *American Journal of Pathology*, **163**, 827–32.

Ewart-Toland, A., Briassouli, P., de Koning, J. P. *et al.* (2003). Identification of *Stk6/STK15* as a candidate low-penetrance tumor-susceptibility gene in mouse and human. *Nature Genetics*, **34**, 403–12.

Ewart-Toland, A., Dai, Q., Gao, Y. T. *et al.* (2005). Aurora-A/STK15 T+91A is a general low penetrance cancer susceptibility gene: a meta-analysis of multiple cancer types. *Carcinogenesis*, **26**, 1368–73.

Farrington, S. M., Lin-Goerke, J., Ling, J. *et al.* (1998). Systematic analysis of h*MSH2* and h*MLH1* in young colon cancer patients and controls. *American Journal of Human Genetics*, **63**, 749–59.

Farrington, S. M., McKinley, A. J., Carothers, A. D. *et al.* (2002). Evidence for an age-related influence of microsatellite instability on colorectal cancer survival. *International Journal of Cancer*, **98**, 844–50.

Farrington, S. M., Tenesa, A., Barnetson, R. *et al.* (2005). Germline susceptibility to colorectal cancer due to Base-Excision Repair Gene Defects. *American Journal of Human Genetics*, **77**, 112–19.

Fearon, E. R., Cho, K. R., Nigro, J. M. *et al.* (1990). Identification of a chromosome 18q gene that is altered in colorectal cancers. *Science*, **247**, 49–56.

Fishel, R. (2001). The selection for mismatch repair defects in hereditary nonpolyposis colorectal cancer: revising the mutator hypothesis. *Cancer Research*, **61**, 7369–74.

Fishel, R., Lescoe, M. K., Rao, M. R. S. *et al.* (1993). The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell*, **75**, 1027–38.

Fodde, R., Kuipers, J., Rosenberg, C. *et al.* (2001). Mutations in the *APC* tumour suppressor gene cause chromosomal instability. *Nature Cell Biology*, **3**, 433–8.

Fuchs, C. S., Giovannucci, E. L., Colditz, G. A. *et al.* (1999). Dietary fiber and the risk of colorectal cancer and adenoma in women. *New England Journal of Medicine*, **340**, 169–76.

Giacosa, A., Franceschi, S., La Vecchia, C. *et al.* (1999). Energy intake, overweight, physical exercise and colorectal cancer risk. *European Journal of Cancer Prevention*, **8** Supplement 1, S53–S60.

Giardello, F. M., Welsh, S. B., Hamilton, S. R. *et al.* (1987). Increased risk of cancer in the Peutz-Jeghers syndrome. *New England Journal of Medicine*, **316**, 1511–14.

Giardiello, F. M., Hamilton, S. R., Krush, A. J. *et al.* (1993). Treatment of colonic and rectal adenomas with sulindac in familial adenomatous polyposis. *New England Journal of Medicine*, **328**, 1313–16.

Giovannucci, E. (2001). An updated review of the epidemiological evidence that cigarette smoking increases risk of colorectal cancer. *Cancer Epidemiology Biomarkers & Prevention*, **10**, 725–31.

Globocan. (2000). *Cancer Incidence, Mortality and prevalence worldwide*. IARC database, 2000. (www.who.int)

Grady, W. M. (2004). Genomic instability and colon cancer. *Cancer Metastasis Reviews*, **23**, 11–27.

Grieu, F., Malaney, S., Ward, R. *et al.* (2003). Lack of association between *CCND1* G870A polymorphism and the risk of breast and colorectal cancers. *Anticancer Research*, **23**, 4257–9.

Grilley, M., Holmes, J., Yashar, B. and Modrich, P. (1990). Mechanisms of DNA-mismatch correction. *Mutation Research*, **26**, 253–67.

Groden, J., Gelbert, L., Thliveris, A. *et al.* (1993). Mutational analysis of patients with adenomatous polyposis: identical inactivating mutations in unrelated individuals. *American Journal of Human Genetics*, **52**, 263–72.

Gryfe, R., Kim, H., Hsieh, E. T. K. *et al.* (2000). Tumor microsatellite instability and clinical outcome in young patients with colorectal cancer. *New England Journal of Medicine*, **342**, 69–77.

Gu, Y., Parker, A., Wilson, T. M. *et al.* (2002). Human *MutY* homolog, a DNA glycosylase involved in base excision repair, physically and functionally interacts with mismatch repair proteins human *MutS* homolog 2/human *MutS* homolog 6. *Journal of Biological Chemistry*, **277**, 11135–42.

Guo, G., Wang, W. and Bradley, A. (2004). Mismatch repair genes identified using genetic screens in *Blm*-deficient embryonic stem cells. *Nature*, **429**, 891–5.

Hackman, P., Tannergard, P., Osei-Mensa, S. *et al.* (1997). A compound heterozygote for two MLH1 missense mutations. *Nature Genetics*, **17**, 135–6.

Hahn, S. A., Schutte, M., Hoque, A. T. M. S. *et al.* (1996). *DPC4*, A candidate tumor-suppressor gene at human-chromosome 18q21.1. *Science*, **271**, 350–3.

Hamilton, S. R., Parsons, R. E., Papadopoulos, N. *et al.* (1995). The molecular basis of Turcot's syndrome. *New England Journal of Medicine*, **332**, 839–47.

Hahnloser, D., Petersen, G. M., Rabe, K. *et al.* (2003). The APC E1317Q variant in adenomatous polyps and colorectal cancers. *Cancer Epidemiology Biomarkers & Prevention*, **12**,1023–8.

He, T.-C., Sparks, A. B., Rago, C. *et al.* (1998). Identification of c-MYC as a target of the APC pathway. *Science*, **281**, 1509–12.

Hemminki, A., Markie, D., Tomlinson, I. *et al.* (1998). A serine/threonine kinase gene defective in Peutz-Jeghers syndrome. *Nature*, **391**, 184–7.

Hienonen, T., Laiho, P., Salovaara, R. *et al.* (2003). Little evidence for involvement of *MLH3* in colorectal cancer predisposition. *International Journal of Cancer*, **106**, 292–6.

Horii, A., Han, H.-J., Shimada, M. *et al.* (1994). Frequent replication errors at microsatellite loci in tumours of patients with multiple primary cancers. *Cancer Research*, **54**, 3373–82.

Howe, J. R., Roth, S., Ringold, J. C. *et al.* (1998). Mutations in the *SMAD4/DPC4* gene in juvenile polyposis. *Science*, **280**, 1086–8.

Howe, J. R., Bair, J. L., Sayed, M. G. *et al.* (2001). Germline mutations of the gene encoding bone morphogenetic protein receptor 1A in juvenile polyposis. *Nature Genetics*, **28**,184–7.

Huang, J., Papadopoulos, N., McKinley, A. J. *et al.* (1996). APC mutations in colorectal tumors with mismatch repair deficiency. *Proceedings of the National Academy of Sciences of the USA*, **93**, 9049–54.

IARC, WHO (1997). *Cancer incidence in five continents*. Lyon: IARC Scientific Publications.

Ionov, Y., Peinado, M. A., Malkhosyan, S. *et al.* (1993). Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature*, **363**, 558–61.

Jaeger, E. E., Woodford-Richens, K. L., Lockett, M. *et al.* (2003). An ancestral Ashkenazi haplotype at the *HMPS/CRAC1* locus on 15q13-q14 is associated with hereditary mixed polyposis syndrome. *American Journal of Human Genetics*, **72**, 1261–7.

Jass, J. R., Smyrk, T. C., Stewart, S. M. *et al.* (1994). Pathology of hereditary non-polyposis colorectal cancer. *Anticancer Research*, **14**, 1631–4.

Jenne, D. E., Reimann, H., Nezu, J. *et al.* (1998). Peutz-Jeghers syndrome is caused by mutations in a novel serine threonine kinase. *Nature Genetics*, **18**, 38–44.

Jiricny, J., Hughes, M., Corman, N. *et al.* (1988). A human 200-kDa protein binds selectively to DNA fragments containing GT mismatches. *Proceedings of the National Academy of Sciences of the USA*, **85**, 8860–4.

Joslyn, L., Carlson, M., Thliveris, A. *et al.* (1991). Identification of deletion mutations and three new genes at the familial polyposis locus. *Cell*, **66**, 601–13.

Kaklamani, V. G., Hou, N., Bian, Y. *et al.* (2003). TGFBR1*6A and cancer risk: a meta-analysis of seven case-control studies. *Journal of Clinical Oncology*, **21**, 3236–43.

Kaplan, K. B., Burds, A. A., Swedlow, J. R. *et al.* (2001). A role for the Adenomatous Polyposis Coli protein in chromosome segregation. *Nature Cell Biology*, **3**, 429–32.

Kemp, Z., Thirlwell, C., Sieber, O. *et al.* (2004). An update on the genetics of colorectal cancer. *Human Molecular Genetics*, **13**, 177–85.

Kinzler, K. W., Nilbert, M. C., Su, L.-K. *et al.* (1991). Identification of FAP locus genes from chomosome 5q21. *Science*, **253**, 661–4.

Kilpivaara, O., Laiho, P., Aaltonen, L. A. and Nevanlinna, H. (2003). *CHEK2* 1100delC and colorectal cancer. *Journal of Medical Genetics*, **40**, e110.

Koi, M., Umar, A., Chauhan, D. P. *et al.* (1994). Human chromosome 3 corrects mismatch repair deficiency and microsatellite instability and reduces N-methyl-N′-nitro-N-nitrosoguanidine tolerance in colon tumor cells with homozygous hMLH1 mutation. *Cancer Research*, **54**, 4308–14.

Kolodner, R., Hall, N., Lipford, J. *et al.* (1994). Structure of the human *MSH2* locus and analysis of two Muir–Torre kindreds for MSH2 mutations. *Genomics*, **24**, 516–26.

Kong, S., Wei, Q., Amos, C. I. *et al.* (2001). Cyclin D1 polymorphism and increased risk of colorectal cancer at young age. *Journal of the National Cancer Institute*, **93**, 1106–8.

La Vecchia, C., Negri, E., Decarli, A. *et al.* (1997). Diabetes mellitus and colorectal cancer risk. *Cancer Epidemiology Biomarkers & Prevention*, **6**, 1007–10.

Laiho, P., Hienonen, T., Karhu, A. *et al.* (2003). Genome-wide allelotyping of 104 Finnish colorectal cancers reveals an excess of allelic imbalance in chromosome 20q in familial cases. *Oncogene*, **22**, 2206–14.

Laken, S. J., Petersen, G. M., Gruber, S. B. *et al.* (1997). Familial colorectal cancer in Ashkenazim due to a hypermutable tract in APC. *Nature Genetics*, **17**, 79–83.

Lamlum, H., Al Tassan, N., Jaeger, E. *et al.* (2000). Germline *APC* variants in patients with multiple colorectal adenomas, with evidence for the particular importance of E1317Q. *Human Molecular Genetics*, **9**, 2215–21.

Lammi, L., Arte, S., Somer, M. *et al.* (2004). Mutations in *AXIN2* cause familial tooth agenesis and predispose to colorectal cancer. *American Journal of Human Genetics*, **74**, 1043–50.

Le Marchand, L., Seifried, A., Lum-Jones, A. *et al.* (2003). Association of the cyclin D1 A870G polymorphism with advanced colorectal cancer. *Journal of American Medical Association*, **290**, 2843–8.

Leach, F. S., Nicolaides, N. C., Papadopolous, N. *et al.* (1993). Mutations of a *MutS* homolog in hereditary non-polyposis colorectal cancer. *Cell*, **75**, 1215–25.

Leppert, M., Dobbs, M., Scambler, P. *et al.* (1987). The gene for familial polyposis maps to the long arm of chromosome 5. *Science*, **238**, 1411–13.

Li, D. M. and Sun, H. (1997). TEP1, encoded by a candidate tumor suppressor locus, is a novel protein tyrosine phosphatase regulated by transforming growth factor beta. *Cancer Research*, **57**, 2124–9.

Liaw, D., Marsh, D. J., Li, J. *et al.* (1997). Germline mutations of the PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome. *Nature Genetics*, **16**, 64–7.

Lichtenstein, P., Holm, N. V., Verkasalo, P. K. *et al.* (2000). Environmental and heritable factors in the causation of cancer–analyses of cohorts of twins from Sweden, Denmark, and Finland. *New England Journal of Medicine*, **343**, 78–85.

Lindblom, A., Tannergard, P., Werelius, B. and Nordenskjold, M. (1993). Genetic mapping of a second locus predisposing to hereditary non-polyposis colon cancer. *Nature Genetics*, **5**, 279–82.

Lipkin, S. M., Wang, V., Jacoby, R. *et al.* (2000). *MLH3*: a DNA mismatch repair gene associated with mammalian microsatellite instability. *Nature Genetics*, **24**, 27–35.

Lipkin, S. M., Rozek, L. S., Rennert, G. *et al.* (2003). The MLH1 D132H variant is associated with susceptibility to sporadic colorectal cancer. *Nature Genetics*, **36**, 694–9.

Liu, B., Farrington, S. M., Petersen, G. M. *et al.* (1995). Genetic instability occurs in the majority of young patients with colorectal cancer. *Nature Medicine*, **2**, 348–52.

Lothe, R. A., Peltomaki, P., Meling, G. I. *et al.* (1993). Genomic instability in colorectal cancer: relationship to clinicopathological variables and family history. *Cancer Research*, **53**, 5849–52.

Lu, S.-L., Kawabata, M., Imamura, T. *et al.* (1998). HNPCC associated with germline mutation in the TGF-beta type II receptor gene. *Nature Genetics*, **19**, 17–18.

Lynch, H. T., Kimberling, W. J., Albano, W. A. *et al.* (1985). Hereditary nonpolyposis colorectal cancer (Lynch Syndromes 1 and 2). 1. Clinical description of resource. *Cancer*, **56**, 934–8.

MacPhee, M., Chepenik, K. P., Liddell, R. A. *et al.* (1995). The secretory phospholipase A2 gene is a candidate for the Mom1 locus, a major modifier of Apc(Min)-induced intestinal neoplasia. *Cell*, **81**, 957–66.

Mandl, M., Paffenholz, R., Friedl, W. *et al.* (1994). Frequency of common and novel inactivating *APC* mutations in 202 families with familial adenomatous polyposis. *Human Molecular Genetics*, **3**, 181–4.

Markowitz, S., Wang, J., Myeroff, L. *et al.* (1995). Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. *Science*, **268**, 1336–8.

Matsumine, A., Ogai, A., Senda, T. *et al.* (1996). Binding of APC to the human homolog of the Drosophila discs large tumour suppressor protein. *Science*, **272**, 1020–3.

Mazurek, A., Berardini, M. and Fishel, R. (2002). Activation of human *MutS* homologs by 8-oxo-guanine DNA damage. *Journal of Biological Chemistry*, **277**, 8260–6.

Meijers-Heijboer, H., van den Ouweland, A., Klijn, J. *et al.* (2002). Low-penetrance susceptibility to breast cancer due to *CHEK2*(*)1100delC in non-carriers of *BRCA1* or *BRCA2* mutations. *Nature Genetics*, **31**, 55–9.

Miyaki, M., Konishi, M., Tanaka, K. *et al.* (1997). Germline mutation of *MSH6* as the cause of hereditary nonpolyposis colorectal cancer. *Nature Genetics*, **17**, 271–2.

Moser, A. R., Dove, W. F., Roth, K. A. *et al.* (1992). The Min (Multiple Intestinal Neoplasia) Mutation: Its effect on gut epithelial cell differentiation and interaction with a modifier system. *Journal of Cell Biology*, **116**, 1517–26.

Moser, A. R., Pitot, H. C. and Dove, W. F. (1990). A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse. *Science*, **247**, 322–4.

Munemitsu, S., Souza, B., Muller, O. *et al.* (1994). The *APC* gene product associates with microtubules in vivo and promotes their assembly in vitro. *Cancer Research*, **54**, 3676–81.

Nagase, H., Miyoshi, Y., Horii, A. *et al.* (1992). Correlation between the location of germ-line mutations in the *APC* gene and the number of colorectal polyps in familial adenomatous polyposis patients. *Cancer Research*, **52**, 4055–7.

Nagase, H. and Nakamura, Y. (1993). Mutations of the *APC* (adenomatous polyposis coli) gene. *Human Mutation*, **2**, 425–34.

Nathke, I.S., Adams, C.L., Polakis, P. *et al.* (1996). The adenomatous polyposis coli (APC) tumour suppressor protein localizes to plasma membrane sites involved in active cell migration. *Journal of Cell Biology*, **134**, 165–80.

Nelen, M. R., Padberg, G. W., Peeters, E. A. *et al.* (1996). Localization of the gene for Cowden disease to chromosome 10q22–23. *Nature Genetics*, **13**, 114–16.

Nicolaides, N. C., Papadopoulos, N., Wei, Y.-F. *et al.* (1994). Mutations of two *PMS* homologues in hereditary non-polyposis colon cancer. *Nature*, **371**, 75–80.

Ohtsubo, T., Nishioka, K., Imaiso, Y. *et al.* (2000). Identification of human *MutY* homolog (hMYH) as a repair enzyme for 2-hydroxyadenine in DNA and detection of multiple forms of hMYH located in nuclei and mitochondria. *Nucleic Acids Research*, **28**, 1355–64.

Olschwang, S., Serova-Sinilnikova, O. M., Lenoir, G. M. and Thomas, G. (1998). *PTEN* germ-line mutations in juvenile polyposis coli. *Nature Genetics*, **18**, 12–14.

Ouyang, H., Furukawa, T., Abe, T. *et al.* (1998). The *BAX* gene, the promoter of apoptosis, is mutated in genetically unstable cancers of the colorectum, stomach, and endometrium. *Clinical Cancer Research*, **4**, 1071–4.

Papadopolous, N., Nicolaides, N. C., Wei, Y.-F. *et al.* (1994). Mutation of a *mutL* homolog in hereditary colon cancer. *Science*, **263**, 1625–9.

Parker, A., Gu, Y., Mahoney, W. *et al.* (2001). Human homolog of the MutY repair protein (hMYH) physically interacts with proteins involved in long patch DNA base excision repair. *Journal of Biological Chemistry*, **276**, 5547–55.

Parkin, D. M. (1992). Studies of cancer in migrant populations: methods and interpretation. *Reviews in Epidemiology Sante Publique*, **40**, 410–24.

Paul, P., Letteboer, T., Gelbert, L. *et al.* (1993). Identical *APC* exon 15 mutations result in a variable phenotype in familial adenomatous polyposis. *Human Molecular Genetics*, **2**, 925–31.

Peltomaki, L., Aaltonen, L. A., Sistonen, P. *et al.* (1993). Genetic mapping of a locus predisposing to human colorectal cancer. *Science*, **260**, 810–12.

Peterlongo, P., Howe, L. R., Radice, P. *et al.* (2005). Germline mutations of *AXIN2* are not associated with non-syndromic colorectal cancer. *Human Mutation*, **25**, 498–500.

Polakis, P. (1999). The oncogenic activation of beta-catenin. *Current Opinion in Genetics & Development*, **9**, 15–21.

Potter J. D. (1999). Colorectal cancer: molecules and populations. *Journal of the National Cancer Institute*, **91**, 916–32.

Ricciardone, M. D., Oezcelik, T., Cevher, B. *et al.* (1999). Human MLH1 deficiency predisposes to hematological malignancy and neurofibromatosis type 1. *Cancer Research*, **59**, 290–3.

Rijnkels, J. M., Hollanders, V. M., Woutersen, R. A. *et al.* (1998). Absence of an inhibitory effect of a vegetables-fruit mixture on the initiation and promotion phases of azoxymethane-induced colorectal carcinogenesis in rats fed low- or high-fat diets. *Nutrition & Cancer*, **30**, 124–9.

Sakatani, T., Kaneda, A., Iacobuzio-Donahue, C. A. *et al.* (2005). Loss of imprinting of *Igf2* alters intestinal maturation and tumorigenesis in mice. *Science*, **307**, 1976–8.

Sampson, J. R., Dolwani, S., Jones, S. *et al.* (2003). Autosomal recessive colorectal adenomatous polyposis due to inherited mutations of MYH. *Lancet*, **362**, 39–41.

Schatzkin, A., Lanza, E., Corle, D. *et al.* (2000). Lack of effect of a low-fat, high-fiber diet on the recurrence of colorectal adenomas. Polyp Prevention Trial Study Group. *New England Journal of Medicine*, **342**, 1149–55.

Sharp, L. and Little, J. (2004). Polymorphisms in genes involved in folate metabolism and colorectal neoplasia: a HuGE review. *American Journal of Epidemiology*, **159**, 423–43.

Shimodaira, H., Filosi, N., Shibata, H. *et al.* (1998). Functional analysis of human *MLH1* mutations in *Saccharomyces cerevisiae*. *Nature Genetics*, **19**, 384–9.

Sieber, O. M., Lipton, L., Crabtree, M. *et al.* (2003). Multiple colorectal adenomas, classic adenomatous polyposis, and germ-line mutations in MYH. *New England Journal of Medicine*, **348**, 791–9.

Sirard, C., de la Pompa, J. L., Elia, A. *et al.* (1998). The tumor suppressor gene *Smad4/Dpc4* is required for gastrulation and later for anterior development of the mouse embryo. *Genes and Development*, **12**, 107–19.

Slattery, M. L., Samowtiz, W., Ballard, L. *et al.* (2001). A molecular variant of the *APC* gene at codon 1822: its association with diet, lifestyle, and risk of colon cancer. *Cancer Research*, **61**, 1000–4.

Slattery, M. L., Samowtiz, W., Ma, K. *et al.* (2004). CYP1A1, cigarette smoking, and colon and rectal cancer. *American Journal of Epidemiology*, **160**, 842–52.

Slattery, M. L., Sweeney, C., Murtaugh, M. *et al.* (2005). Associations between ApoE genotype and colon and rectal cancer. *Carcinogenesis*, (In press (E-pub, April)).

Smith, K.J., Johnson, K.A., Bryan, T.M. *et al.* (1993). The APC gene product in normal and tumour cells. *Proceedings of the National Academy of Sciences of the USA*, **90**, 2846–50.

Smith, K.J., Levy, D.B., Maupin, P. *et al.* (1994). Wild-type but not mutant APC associates with microtubule cytoskeleton. *Cancer Research*, **54**, 3672–5.

Stark, L.A. and Dunlop, M.G. (2005). Nucleolar sequestration of Rel A (p65) regulates NF-κB driven transcription and apoptosis. *Molecular and Cellular Biology*, **25**, 5985–6004.

Steinbach, G., Lynch, P.M., Phillips, R.K. *et al.* (2000). The effect of celecoxib, a cyclooxygenase-2 inhibitor, in familial adenomatous polyposis. *New England Journal of Medicine*, **342**, 1946–52.

Strand, M., Prolla, T.A., Liskay, R.M. *et al.* (1993). Destabilization of tracts of simple repetitive DNA in yeasts by mutations affecting DNA mismatch repair. *Nature*, **365**, 274–6.

Su, L.-K., Burrell, M., Hill, D.E. *et al.* (1995). APC binds to the novel protein EB1. *Cancer Research*, **55**, 2972–7.

Su, L.-K., Johnson, K.A., Smith, K.J. *et al.* (1993). Association between wild-type and mutant *APC* gene products. *Cancer Research*, **53**, 2728–31.

Suter, C.M., Martin, D.I. and Ward, R.L. (2004). Germline epimutation of *MLH1* in individuals with multiple cancers. *Nature Genetics*, **36**, 497–501.

Thibodeau, S.N., Bren, G. and Schaid, D. (1993). Microsatellite instability in cancer of the proximal colon. *Science*, **260**, 816–19.

Thomas, H.J.W., Whitelaw, S.C., Cottrell, S.E. *et al.* (1996). Genetic mapping of hereditary mixed polyposis syndrome to chromosome 6q. *American Journal of Human Genetics*, **58**, 770–6.

Tomlinson, I.P.M., Neale, K., Talbot, I.C. *et al.* (1996). A modifying locus for familial adenomatous polyposis may be present on chromosome 1p35-p36. *Journal of Medical Genetics*, **33**, 268–73.

Toyota, M., Ahuja, N., Ohe-Toyota, M. *et al.* (1999). CpG island methylator phenotype in colorectal cancer. *Proceedings of the National Academy of Sciences of the USA*, **96**, 8681–6.

Trojan, J., Zeuzem, S., Randolph, A. *et al.* (2002). Functional analysis of h*MLH1* variants and HNPCC-related mutations using a human expression system. *Gastroenterology*, **122**, 211–19.

Van der Houven van Oordt, C.W., Smits, R., Schouten, T.G. *et al.* (1999). The genetic background modifies the spontaneous and X-ray-induced tumor spectrum in the Apc1638N mouse model. *Genes, Chromosomes & Cancer*, **24**, 191–8.

van Kranen, H.J., van Iersel, P.W., Rijnkels, J.M. *et al.* (1998). Effects of dietary fat and a vegetable-fruit mixture on the development of intestinal neoplasia in the *ApcMin* mouse. *Carcinogenesis*, **19**, 1597–601.

Vasen, H.F.A., Wijnen, J.T., Menko, F.H. *et al.* (1996). Cancer risk in families with hereditary nonpolyposis colorectal cancer diagnosed by mutation analysis. *Gastroenterology*, **110**, 1020–7.

Venesio, T., Molatore, S., Cattaneo, F. *et al.* (2004). High frequency of *MYH* gene mutations in a subset of patients with familial adenomatous polyposis. *Gastroenterology*, **126**, 1681–5.

Wang, L., Baudhuin, L.M., Boardman, L.A. *et al.* (2004). *MYH* mutations in patients with attenuated and classic polyposis and with young-onset colorectal cancer without polyps. *Gastroenterology*, **127**, 9–16.

Wang, Q., Lasset, C., Desseigne, F. *et al.* (1999). Neurofibromatosis and early onset of cancers in h*MLH1*-deficient children. *Cancer Research*, **59**, 294–7.

Watanabe, A., Ikejima, M., Suzuki, N. *et al.* (1996). Genomic organization and expression of the human *MSH3* gene. *Genomics*, **31**, 311–18.

Wei, K., Kucherlapati, R. and Edelmann, W. (2002). Mouse models for human DNA mismatch-repair gene defects. *Trends in Molecular Medicine*, **8**, 346–53.

Wiebauer, K.M. and Jiricny, J. (1990). Mismatch-specific thymine DNA glysosylase and DNA polymerase B mediate the correction of G.T mispairs in nuclear extracts from human cells. *Proceedings of the National Academy of Sciences of the USA*, **87**, 5842–5.

Wijnen, J.T., de Leeuw, W., Vasen, H. *et al.* (1999). Familial endometrial cancer in female carriers of *MSH6* germline mutations. *Nature Genetics*, **23**, 142–4.

World Cancer Research Fund Panel (Chair: J.D. Potter). (2001). *Diet, nutrition, and the prevention of cancer*. (www.wcrf.org.uk).

Yang, G., Scherer, S.J., Shell, S.S. *et al.* (2004). Dominant effects of an *Msh6* missense mutation on DNA repair and cancer susceptibility. *Cancer Cell*, **6**, 139–50.

Yang, K., Edelmann, W., Fan, K. *et al.* (1998). Dietary modulation of carcinoma development in a mouse model for human familial adenomatous polyposis. *Cancer Research*, **58**, 5713–17.

Zhang, H., Richards, B., Wilson, T. *et al.* (1999). Apoptosis induced by overexpression of h*MSH2* or h*MLH1*. *Cancer Research*, **59**, 3021–7.

18

Genetics of autoimmune disease

John I. Bell and Lars Fugger

Introduction

The autoimmune diseases are disorders where the immune system erroneously targets self-antigens leading to organ specific or systemic tissue damage (Goodnow *et al.*, 2005). These aberrant responses are the results of a breach in tolerance characterized by either cellular or humoral autoreactivity. Examples of autoimmunity occur in most organ systems and include autoimmune responses to the β-cell in the pancreas (Type 1 diabetes), to myelin in the central nervous system (multiple sclerosis), to hepatocytes (chronic active hepatitis), to thyrocytes (Graves' disease or autoimmune thyroiditis), skin (psoriasis or vitiligo), renal basement membrane (Goodpasture's disease), joints (rheumatoid arthritis) or gastrointestinal mucosa (ulcerative colitis or Crohn's disease). Generalized autoimmune diseases such as systemic lupus erythematosus, and system sclerosis also affect multiple organs simultaneously.

Background: genes and environment

The characterization of genetic determinants in disease has been transformed with the availability of the tools and technologies which have emerged from sequencing and characterizing variation in the human genome over the past 20 years. This has allowed the systematic characterization of the genome and the analysis of variation in genetic material in individuals and in families with a range of complex common diseases. The autoimmune diseases represent typical complex diseases with substantial genetic and environmental components (Roderick and Navajas, 2003). The characterization of the genetic determinants responsible for these diseases began 20 years before genome sequencing began (Brewerton *et al.*, 1973) and more recently the characterization of the multiple genetic determinants responsible for these diseases has proved to be a paradigm for those searching for diseases genes in other common conditions (Rioux and Abbas, 2005). The significant advantage for those studying the genetics of autoimmune diseases is that the main susceptibility loci lie within the major histocompatibility complex (MHC) (see Chapter 9, this volume and also Kelley *et al.*, 2005). This complex encodes the HLA (human leukocyte antigen) loci and was of significant interest well before DNA-based genotyping became available, because of the extensive polymorphisms in HLA molecules that could be typed using antisera from multiparous women. HLA serological typing therefore allowed the earliest characterization of disease gene susceptibility loci. As a result of the work on HLA typing that occurred in the late 1960s and 1970s, many genetic determinants for autoimmune disease were attributed to variation within the HLA through genetic association studies (Svejgaard and Ryder, 1994). Many of the important lessons learned through this period about the contribution of linkage disequilibrium, genetic mapping and the importance of multiple alleles of a single locus have contributed greatly

to more recent studies using DNA-based methodologies in a range of common diseases. Importantly, *HLA* alleles when characterized in detail at a genetic level have revealed mechanisms for autoimmune disease pathogenesis which have greatly facilitated our understanding of these diseases (Todd *et al.*, 1988). These loci have also proved to date to be by far the most important genetic determinants in most forms of autoimmunity and although other non-HLA loci have been characterized, these have usually proved to be much less significant determinants (Julier *et al.*, 1991; Ueda *et al.*, 2003).

Autoimmunity, therefore, presents an opportunity to understand the contribution of at least one highly polymorphic locus in detail in many diseases. In addition, however, it has proved in some diseases to be a model system for the characterization of non-HLA loci. As a result, our understanding of the genetic basis of a disease such as inflammatory bowel disease has revolutionized our understanding of the pathogenesis of this disease (see Chapter 20, this volume and also Mathew and Lewis, 2004). Few complex diseases have proved to be as tractable genetically as autoimmune disease.

Despite the relative success in identifying disease genes in a range of autoimmune diseases and a recognition that the major genes associated with many autoimmune diseases (HLA) are normally involved in antigen presentation to T-lymphocytes, it has been much more difficult to clarify the role of environmental factors in mediating disease initiation in any of these disorders. This environmental component of a disease has been studied both at a biological and population level but, to date no specific environmental factor has been related to causality in these disorders. Information abounds about specific peptides which are capable of stimulating T-cell responses when bound to the appropriate HLA susceptibility antigens. A collagen peptide binds strongly to HLA DR4 (Dessen *et al.*, 1997), the susceptibility determinant in rheumatoid arthritis, while a myelin basic protein peptide binds to alleles associated with susceptibility to multiple sclerosis (Madsen *et al.*, 1999). Despite these leads, however, it has not been possible to demonstrate a direct relationship between these peptides and the pathophysiology of these diseases although molecular mimicry has been suggested in many circumstances where peptides from pathogens share some sequence or structural homology (Lang *et al.*, 2002; Wekerle and Hohlfeld, 2003).

At a population level there has been much speculation around the geographical and ethnic distribution of these diseases. For example, multiple sclerosis becomes more prevalent as one moves to latitudes distant from the equator (Sadovnick and Ebers, 1993) and rheumatoid arthritis has very high prevalences in some aboriginal populations in Canada and Africa (Harney and Wordsworth, 2002). Exposure to environmental risk factors has been described; for example, a role for vitamin D has been suggested in the pathogenesis of multiple sclerosis (Ebers *et al.*, 2004) and exposure to cow's milk has been suggested to be involved in susceptibility to Type 1 diabetes (Virtanen *et al.*, 2000). Despite this extensive speculation, however, no concrete and reproducible data has yet been provided to demonstrate the role of an environmental exposure in disease and, in particular, no intervention study has been undertaken that can clearly assign a role for these various factors. The tools provided by genetics have, however, provided a range of experimental systems that have allowed the exploration of potential mechanisms for autoimmune disease and one of these is described below. In pure observational epidemiology terms, the rising incidence of disease such as Type 1 diabetes (Gillespie *et al.*, 2004) argues strongly for an important environmental determinant which is changing in frequency. Little else could explain the relatively dramatic changes in incidence of these disorders.

The most persuasive theories of environmental exposure and risk of autoimmune disease relate to an individual's exposure to infectious pathogens. The suggestion that initially a normal immune response to an infectious agent might initiate an

organ-specific immune reaction has been supported by clinical syndromes akin to major autoimmune diseases after infection with Coxsackie virus (leading to β-cell destruction and Type 1 diabetes) (Green *et al.*, 2004), inflammatory joint diseases after exposure to a range of viral or microbial infections including Reiter's syndrome and peripheral arthritis (Moreland and Koopman, 1992). In addition, much evidence has now accumulated for molecular mimicry, where the response to a HLA-restricted peptide from an infectious pathogen leads to a T-cell response which cross-reacts with a self-antigen. Despite extensive efforts to clarify the role of infectious agents in mediating autoimmunity, this has not generated conclusive evidence in any single disease, although there remain strong suspicions that these agents play a role in the initiation of these disorders.

HLA: the major genetic locus in autoimmunity

The first recognition of a role for a specific gene for determining autoimmune susceptibility came from the identification in 1973 that the vast majority of patients with ankylosing spondilitis, a disorder of axial joints, possessed *HLA B-27* (Brewerton *et al.*, 1973). This observation was made utilizing HLA antisera rather than DNA-based genetic methodology and has led to the transformation in our understanding of the genetics of autoimmunity and of the functional role of HLA antigens themselves. This initial report was shortly followed by similar observations in a range of other autoimmune diseases. In the case of diabetes, investigators identified a role for *HLA B8* in the juvenile form of disease, distinguishing a subtype of diabetes, now known as Type 1 diabetes, from other more common adult forms of the disease (Cudworth and Woodrow, 1976). As the tools for typing HLA variants became better and incorporated both cellular and humoral reagents, it became clear that the exact nature of susceptibility within the MHC was extremely complex. The relationship between disease and HLA is covered elsewhere in this volume. The exploration of determinants of disease lying within the MHC has revealed the opportunities and challenges associated with studying regions of strong linkage disequilibrium and the problems associated with identifying precise genetic variants mediating disease when multiple loci within the same region may contribute to pathogenesis. These lessons have been relearned repeatedly in other regions of the genome associated with disease genes over the past 15 years. Importantly, allelic variants within the MHC can contribute both to susceptibility and to resistance to autoimmune disease, and alleles at these loci have also been demonstrated to confer resistance and susceptibility to a range of infectious pathogens (Segal and Hill, 2003) including major pathogens which drive significant selection in human populations such as HIV and malaria. This strongly suggests that the driving evolutionary pressure that has created and fixed such extensive polymorphism in the population has been exposure to infectious pathogens.

Molecular dissection at a DNA level of allelic variants within the MHC has ultimately delivered a better understanding of the precise alleles associated with some autoimmune diseases than was ever achieved using antisera. Notable among these are the role of trans-heterodimers in mediating celiac disease (Sollid, 2002), and the role of HLA-DR and DQ motifs in mediating susceptibility to Type 1 diabetes (Todd *et al.*, 1987) and rheumatoid arthritis (Wordsworth *et al.*, 1989).

Interestingly, comparisons using linear DNA sequence ultimately provide only part of the explanation of HLA disease susceptibility. Sequences of DNA confer their ultimate contribution to risk through protein structure and function and the definitive approach to understanding the contribution to disease susceptibility from multiple alleles lies in an understanding of the similarities and differences in three dimensions within,

for example, the HLA and peptide binding groove. Systematic analysis of a range of HLA DQ crystal structures, for example, has revealed an important interplay between the volume of the P6 pocket and the specificity of the P9 pocket in determining protection from Type 1 diabetes, implying that an expanded peptide repertoire is important in producing such protection (Siebold *et al.*, 2004).

It is humbling to recognize that, after 30 years of investigation, the precise nature of HLA-encoded susceptibility to many autoimmune diseases is still not fully refined and that the mechanism by which these alleles confer their susceptibility, in particular their interaction with environmental stimuli, has not yet been demonstrated with clarity in any single autoimmune disease. This indicates the challenge ahead in determining a precise role for less important genetic factors in these and other complex diseases.

Genetic approaches to defining non-HLA susceptibility genes

The basis for understanding genetic contributions to autoimmunity was therefore laid around HLA in the era before tools became available for studying other DNA variants and their contribution to disease. The availability of a large set of highly variable polymorphic markers in the form of microsatellites allowed systematic analyses of the human genome sufficient to characterize linkage in large numbers of families (Dib *et al.*, 1996). The autoimmune diseases, perhaps because of the existing strong understanding of HLA genetic determinants, were among the first to utilize this technology to identify non-HLA susceptibility loci. Papers demonstrating areas of the genome that contained evidence suggesting linkage to Type 1 diabetes appeared first in 1994 (Davies *et al.*, 1994; Hashimoto *et al.*, 1994). Experience has demonstrated that such linkage data utilizing relatively small numbers of families is not particularly robust (Concannon *et al.*, 1998), although the accumulation of many studies and many family sets in Type 1 diabetes has revealed some consistency in genetic linkage beyond the known locus in HLA (Wicker *et al.*, 2005). Interestingly, the two other loci that have been clearly and reproducibly characterized as having a role in Type 1 diabetes are the insulin locus (*INS*) and *CTLA-4* and both were characterized as candidate loci rather than having been identified from genome scans.

The region of linkage disequilibrium associated with Type 1 diabetes on chromosome 11p extended from the insulin locus through IGF2 encompassing the *INS* (Julier *et al.*, 1991). It appears after extensive analysis that the primary genetic variant leading to this association lies within the VNTR upstream of the insulin locus (Lucassen *et al.*, 1995). The contribution of this genetic variation can be seen using family-based association and linkage. A second non-*HLA* locus in Type 1 diabetes has also been described, again based on association data. The biological evidence for a role of *CTLA-4* is now accumulating (Ueda *et al.*, 2003) and it would appear that coding sequence variation in this gene, responsible for a cell surface molecule on lymphocytes, may contribute to Type 1 diabetes susceptibility. This disease has proved to be a model for genetic studies in autoimmunity and it is revealing that the early promise of genome wide linkage studies has now given way to an association-based strategy either at the candidate gene locus or, as is currently being undertaken, on a larger genome-wide non-hypothesis-driven basis.

Many of the lessons learned from genetic studies in Type 1 diabetes have now proved to be useful in the context of other autoimmune diseases which have been scanned using genome-wide linkage approaches in more recent years. For example, a large-scale study of autoimmune thyroid disease started with 1192 affected sibling pairs, substantially more than the first and second generation Type 1 diabetes scans had utilized a decade before (Taylor *et al.*, 2006). More appropriately powered, this study has revealed interesting regions of linkage around the genome (chromosomes 18, 11, 2). The genetic linkage study in autoimmune

thyroid disease has provided additional insights into the relationship between autoimmune thyroiditis and Graves' disease, the two major clinical forms of the disease. These disorders are distinguished in part by their clinical phenotype and also by their histological characteristics. Analysis of cosegregation of these two clinical phenotypes in the same family and clinical uncertainty about shared features of both subtypes of autoimmune thyroid disease has led to a view that these two disorders are in some ways related. Genome-wide linkage analysis has revealed, however, that where Graves' disease has a relatively strong genetic signal with multiple loci having been identified as leading to susceptibility, autoimmune thyroiditis has a distinct genetic component and the loci identified in this large-scale study do not, in any significant way, overlap. This study therefore suggests that, at the level of genetics and their associated biochemical pathways, these two diseases appear to be aetiologically different. Another interesting observation from this relatively recent genome scan relates to the overlap between general autoimmune susceptibility genes and those associated with a particular tissue target. Analysis of the genetic determinants uncovered in a range of autoimmune diseases indicates that there are many regions that are shared between different autoimmune diseases and others that appear to be relatively disease specific. Although care must be taken not to overinterpret this data as a correlation may exist based on chance alone given the very large number of loci that have now been mapped using whole genome methodology, it would still appear likely that many of these linkages shared between diseases may reflect common autoimmune genes that increase an individual's liability to the loss of self-tolerance (Smyth *et al.*, 2004).

Perhaps the best example of genome-wide linkage followed by positional cloning leading to new insights into the pathogenesis of autoimmune diseases can be found in the characterization of inflammatory bowel disease (IBD). This disorder has several recognized clinical subtypes — Crohn's disease and ulcerative colitis — and the twin concordance data has indicated that the genetic susceptibility in these diseases might be considerably stronger than in other common diseases. In addition, the contribution of HLA, which dominates the genetic effects in diseases like Type 1 diabetes and rheumatoid arthritis, is hardly detectable in Crohn's disease and is only a small contributor to the susceptibility of ulcerative colitis (Ahmad *et al.*, 2001). Genome-wide mapping led initially to the identification of a locus on chromosome 16 in Crohn's disease (Hugot *et al.*, 1996) and then to multiple other loci in ulcerative colitis or Crohn's disease in subsequent genome-wide scans (Satsangi *et al.*, 1996). The locus on chromosome 16 was successfully positionally cloned and shown to be the *NOD2* locus (*CARD-15*) (Hugot *et al.*, 2001) (Chapter 20). This has led to an insight into the mechanism of this disease. It now appears that *NOD2* acts to constrain the pro-inflammatory response resulting from TOLL-like receptor to stimulation (Watanabe *et al.*, 2004). Mutations in *NOD2* prevent this inhibition from occurring leading to a tendency for inflammation to occur. This important genetically led experiment indicated the role of innate immunity in chronic inflammation and autoimmunity. Further insights into the mechanism of this disease came from the characterization of the DNA variant on chromosome 10q linked to IBD through genome wide linkage. This genetic variant lies within the *DLG5* gene locus, the product of which is the cell scaffold responsible for maintaining epithelial integrity (Stoll *et al.*, 2004). It is possible, therefore, that mutations in this locus may reduce the ability of the epithelial barrier to function properly in the colon. Variants at this locus may interact with those at the *NOD2* locus. Another locus to emerge from genome wide association is that on chromosome 5q31. This region is rich in cytokines and transporters, and positional cloning across this region in IBD patients has revealed two distinct haplotypic variants in the organic cation transporter cluster (Peltekova *et al.*, 2004). Additional evidence has implicated the *NOD-1* locus in

disease pathogenesis (McGovern *et al.*, 2005). Again, genetics has revealed a completely novel mechanism for susceptibility in this disorder also shown to be associated with susceptibility to rheumatoid arthritis. Potential mechanisms by which this might mediate disease susceptibility include the reduction of cation transport, defects in oxygen-mediated pathogen clearing, or impaired fatty acid β oxidation. Transporter defects may also reduce the movement of toxins such as those produced by interluminal bacteria and that may mediate some of the effects seen because of *NOD2* mutations. These, therefore, represent an excellent example of how genetic linkage followed by positional cloning has provided exciting new insights into disease mechanisms which had eluded further identification by other non-genetic approaches to disease characterization.

The outcome from more than ten years' work in genome-wide linkage studies in a large range of autoimmune diseases is that we are in many diseases confident of regions of the genome which contain genetic determinants of disease susceptibility and that in some diseases many of these genetic variants have now been uncovered. This has provided important new insights into the mechanisms of disease and new opportunities for therapeutic interventions. In most cases, however, genetic linkages have not progressed to the identification of individual DNA variants and that may be left to more substantial association studies using large numbers of single nucleotide polymorphisms across regions of interest or at relatively high density around the genome.

Using animal models to understand genetic contributions to autoimmunity

Rodent models of autoimmune disease have formed an important component of the scientific strategy to better understand the pathophysiology of these disorders ever since autoimmune diseases were originally described. Genetic tools and manipulation have in recent years allowed these rodent models to assume a new importance. Three major approaches have been used to study autoimmunity in murine models — genetic linkage, genetic manipulation using transgenics and knockout models of individual genes and systematic mutagenesis strategies.

Mouse strains exist which have an innate liability to develop autoimmune disease which resembles the diseases seen in man. Mouse models of autoimmunity that require antigenic induction of the disease phenotype (e.g. extrinsic allergic encephalomyelitis EAE or collagen-induced arthritis) provide an opportunity to map strain-specific susceptibility determinants. Spontaneous forms of autoimmunity are less artificial, the two most widely utilized being the non-obese diabetic mouse (a model of Type I diabetes) and strains of mice which develop syndromes similar to systemic lupus erythematosus (e.g. NZB/NEW, Ipr). In both these models, extensive efforts have been made to characterize the loci and genes responsible for disease susceptibility. One of the challenges which has emerged from these mapping studies is distinguishing strain-specific variants. Characterizing genomic regions that may contribute to disease susceptibility has been relatively successful but the dissection of these regions and the identification of the precise genetic variant and genetic locus responsible for conferring these susceptibilities is much more difficult. Distinguishing strain-specific variants from those which contribute to disease susceptibility has therefore required functional approaches.

The use of genetic manipulation in the mouse has proved to be an additional important tool in characterizing the mechanisms by which particular genes and their variants might mediate autoimmune susceptibility. The use of genetic manipulation has clarified the role of *AIRE-1* gene mutations in mediating autoimmune susceptibility (Anderson *et al.*, 2002). Models of autoimmunity have been created by introducing human autoimmunity-associated genes into mouse backgrounds and again these provide significant insights into

disease mechanisms. For example, EAE, the mouse model of multiple sclerosis, has provided much useful information about the mechanism of inflammatory damage in the central nervous system. This model has been substantially improved by the introduction of genetic variants known to contribute to multiple sclerosis in man. This includes the DRB1*1501 *HLA* allele on the DR2 haplotype and the T-cell receptor that recognizes the myelin basic protein peptide 85–99 in the context of this allele, which have been expressed together in mice to create a humanized model of the disease (Madsen *et al.*, 1999). These animals develop clinical manifestations more consistent with human multiple sclerosis than other EAE models. Interestingly, this T-cell receptor is also capable of recognizing an Epstein–Barr virus peptide in the context of another HLA molecule on the same DR2 haplotype. This raises the possibility that molecular mimicry might in some way be responsible for the activation of the myelin specific immune response. These humanized mice have also allowed the characterization of individual *HLA* alleles both independently and together and have demonstrated important epistatic effects between *HLA-DR* alleles from the DR2 haplotype (Friese *et al.*, 2006). These humanized murine models of autoimmune disease have greatly enhanced our ability to understand the specific role of genes in mediating autoimmune events.

In other mouse models of autoimmunity, genes known to be associated with human autoimmunity have been introduced into the mouse along with the mutations associated with human disease. The mutations associated with the *NOD-2* locus in inflammatory bowel disease have been ''knocked in'' to mice. This model has allowed the characterization of the role of this variant in mediating IBD (Maeda *et al.*, 2005).

In addition to these targeted approaches for introducing genes into murine systems, it has now become possible to more systematically evaluate the role of genes in mediating autoimmune processes. A range of technologies exist for targeted mutagenesis of individual loci using insertion mutagenesis with viruses that may allow mice to be generated with mutations in every genetic locus in the genome. The use of ethylnitrosourea (ENU), a potent chemical carcinogen to produce point mutations, has been highly productive in identifying novel pathways that might be associated with autoimmune processes (Vinuesa *et al.*, 2005). Animals identified with lymphadenopathy and high titres of anti-DNA antibodies have then been characterized for the underlying genetic mutation responsible for this phenotype and, interestingly, these animals have been shown to have mutations in a ubiquitin ligase-like molecule associated with constitutively high levels of pro-inflammatory cytokines and a severe systemic autoimmune phenotype. These approaches have clearly demonstrated that mouse genetic manipulation can be used successfully to identify totally novel molecules and pathways that, if perturbed, can create autoimmune phenotypes.

Conclusions

Understanding the mechanisms underlying autoimmune diseases remains a major scientific challenge. Although slower than expected, the dissection of these diseases at a genetic level in man has in some cases been enormously successful and the introduction of a new, more powerful technology should allow many other genetic loci to be uncovered in the future. These human genetic studies are greatly strengthened by the ability to introduce genetic variation into mouse models and to study the disease in more detail in this arena. This powerful combination of human genetics and mouse genetics is providing remarkable insights and numerous new therapeutic opportunities for this major class of disease.

REFERENCES

Ahmad, T., Satsangi, J., McGovern, D., Bunce, M. and Jewell, D. P. (2001). Review article: the genetics of

inflammatory bowel disease. *Aliment Pharmacol Ther*, **15**, 731–48.

Anderson, M. S., Venanzi, E. S., Klein, L. *et al.* (2002). Projection of an immunological self shadow within the thymus by the aire protein. *Science*, **298**, 1395–401.

Brewerton, D. A., Hart, F. D., Nicholls, A. *et al.* (1973). Ankylosing spondylitis and HL-A 27. *Lancet*, **1**, 904–7.

Concannon, P., Gogolin-Ewens, K. J., Hinds, D. A. *et al.* (1998). A second-generation screen of the human genome for susceptibility to insulin-dependent diabetes mellitus. *Nat Genet*, **19**, 292–6.

Cudworth, A. G. and Woodrow, J. C. (1976). Genetic susceptibility in diabetes mellitus: analysis of the HLA association. *Br Med J*, **2**, 846–8.

Davies, J. L., Kawaguchi, Y., Bennett, S. T. *et al.* (1994). A genome-wide search for human type 1 diabetes susceptibility genes. *Nature*, **371**, 130–6.

Dessen, A., Lawrence, C. M., Cupo, S., Zaller, D. M. and Wiley, D. C. (1997). X-ray crystal structure of HLA-DR4 (DRA*0101, DRB1*0401) complexed with a peptide from human collagen II. *Immunity*, **7**, 473–81.

Dib, C., Faure, S., Fizames, C. *et al.* (1996). A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature*, **380**, 152–4.

Ebers, G. C., Sadovnick, A. D. and Veith, R. (2004). Vitamin D intake and incidence of multiple sclerosis. *Neurology*, **63**, 939; author reply 939.

Friese, M. A., Montalban, X., Willcox, N. *et al.* (2006). The value of animal models for drug development in multiple sclerosis. *Brain*, **129**, 1940–52.

Gillespie, K. M., Bain, S. C., Barnett, A. H. *et al.* (2004). The rising incidence of childhood type 1 diabetes and reduced contribution of high-risk HLA haplotypes. *Lancet*, **364**, 1699–700.

Goodnow, C. C., Sprent, J., Fazekas de St Groth, B. and Vinuesa, C. G. (2005). Cellular and genetic mechanisms of self tolerance and autoimmunity. *Nature*, **435**, 590–7.

Green, J., Casabonne, D. and Newton, R. (2004). Coxsackie B virus serology and Type 1 diabetes mellitus: a systematic review of published case-control studies. *Diabet Med*, **21**, 507–14.

Harney, S. and Wordsworth, B. P. (2002). Genetic epidemiology of rheumatoid arthritis. *Tissue Antigens*, **60**, 465–73.

Hashimoto, L., Habita, C., Beressi, J. P. *et al.* (1994). Genetic mapping of a susceptibility locus for insulin-dependent diabetes mellitus on chromosome 11q. *Nature*, **371**, 161–4.

Hugot, J. P., Chamaillard, M., Zouali, H. *et al.* (2001). Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature*, **411**, 599–603.

Hugot, J. P., Laurent-Puig, P., Gower-Rousseau, C. *et al.* (1996). Mapping of a susceptibility locus for Crohn's disease on chromosome 16. *Nature*, **379**, 821–3.

Julier, C., Hyer, R. N., Davies, J. *et al.* (1991). Insulin-IGF2 region on chromosome 11p encodes a gene implicated in HLA-DR4-dependent diabetes susceptibility. *Nature*, **354**, 155–9.

Kelley, J., Walter, L. and Trowsdale, J. (2005). Comparative genomics of major histocompatibility complexes. *Immunogenetics*, **56**, 683–95.

Lang, H. L., Jacobsen, H., Ikemizu, S. *et al.* (2002). A functional and structural basis for TCR cross-reactivity in multiple sclerosis. *Nat Immunol*, **3**, 940–3.

Lucassen, A. M., Screaton, G. R., Julier, C. *et al.* (1995). Regulation of insulin gene expression by the IDDM associated, insulin locus haplotype. *Hum Mol Genet*, **4**, 501–6.

Madsen, L. S., Andersson, E. C., Jansson, L. *et al.* (1999). A humanized model for multiple sclerosis using HLA-DR2 and a human T-cell receptor. *Nat Genet*, **23**, 343–7.

Maeda, S., Hsu, L. C., Liu, H. *et al.* (2005). Nod2 mutation in Crohn's disease potentiates NF-kappaB activity and IL-1beta processing. *Science*, **307**, 734–8.

Mathew, C. G. and Lewis, C. M. (2004). Genetics of inflammatory bowel disease: progress and prospects. *Hum Mol Genet*, **13** Spec No 1, R161–8.

McGovern, D. P., Hysi, P., Ahmad, T. *et al.* (2005). Association between a complex insertion/deletion polymorphism in NOD1 (CARD4) and susceptibility to inflammatory bowel disease. *Hum Mol Genet*, **14**, 1245–50.

Moreland, L. W. and Koopman, W. J. (1992). Infection as a cause of reactive arthritis, ankylosing spondylitis and rheumatic fever. *Curr Opin Rheumatol*, **4**, 534–42.

Peltekova, V. D., Wintle, R. F., Rubin, L. A. *et al.* (2004). Functional variants of OCTN cation transporter genes are associated with Crohn disease. *Nat Genet*, **36**, 471–5.

Rioux, J. D. and Abbas, A. K. (2005). Paths to understanding the genetic basis of autoimmune disease. *Nature*, **435**, 584–9.

Roderick, G. K. and Navajas, M. (2003). Genes in new environments: genetics and evolution in biological control. *Nat Rev Genet*, **4**, 889–99.

Sadovnick, A. D. and Ebers, G. C. (1993). Epidemiology of multiple sclerosis: a critical overview. *Can J Neurol Sci*, **20**, 17–29.

Satsangi, J., Parkes, M., Louis, E. *et al.* (1996). Two stage genome-wide search in inflammatory bowel disease provides evidence for susceptibility loci on chromosomes 3, 7 and 12. *Nat Genet*, **14**, 199–202.

Segal, S. and Hill, A. V. (2003). Genetic susceptibility to infectious disease. *Trends Microbiol*, **11**, 445–8.

Siebold, C., Hansen, B. E., Wyer, J. R. *et al.* (2004). Crystal structure of HLA-DQ0602 that protects against type 1 diabetes and confers strong susceptibility to narcolepsy. *Proc Natl Acad Sci USA*, **101**, 1999–2004.

Smyth, D., Cooper, J. D., Collins, J. E. *et al.* (2004). Replication of an association between the lymphoid tyrosine phosphatase locus (LYP/PTPN22) with type 1 diabetes and evidence for its role as a general autoimmunity locus. *Diabetes*, **53**, 3020–3.

Sollid, L. M. (2002). Coeliac disease: dissecting a complex inflammatory disorder. *Nat Rev Immunol*, **2**, 647–55.

Stoll, M., Corneliussen, B., Costello, C. M. *et al.* (2004). Genetic variation in DLG5 is associated with inflammatory bowel disease. *Nat Genet*, **36**, 476–80.

Svejgaard, A. and Ryder, L. P. (1994). HLA and disease associations: detecting the strongest association. *Tissue Antigens*, **43**, 18–27.

Taylor, J. C., Gough, S. C., Hunt, P. J. *et al.* (2006). A genome-wide screen in 1119 relative pairs with autoimmune thyroid disease. *J Clin Endocrinol Metab*, **91**, 646–53.

Todd, J. A., Acha-Orbea, H., Bell, J. I. *et al.* (1988). A molecular basis for MHC class II–associated autoimmunity. *Science*, **240**, 1003–9.

Todd, J. A., Bell, J. I. and McDevitt, H. O. (1987). HLA-DQ beta gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. *Nature*, **329**, 599–604.

Ueda, H., Howson, J. M., Esposito, L. *et al.* (2003). Association of the T-cell regulatory gene *CTLA4* with susceptibility to autoimmune disease. *Nature*, **423**, 506–11.

Vinuesa, C. G., Cook, M. C. Angelacci, C. *et al.* (2005). A RING-type ubiquitin ligase family member required to repress follicular helper T cells and autoimmunity. *Nature*, **435**, 452–8.

Virtanen, S. M., Laara, E., Hypponen, E. *et al.* (2000). Cow's milk consumption, HLA-DQB1 genotype and type 1 diabetes: a nested case-control study of siblings of children with diabetes. Childhood diabetes in Finland study group. *Diabetes*, **49**, 912–17.

Watanabe, T., Kitani, A., Murray, P. J. and Strober, W. (2004). NOD2 is a negative regulator of Toll-like receptor 2-mediated T helper type 1 responses. *Nat Immunol*, **5**, 800–8.

Wekerle, H. and Hohlfeld, R. (2003). Molecular mimicry in multiple sclerosis. *N Engl J Med*, **349**, 185–6.

Wicker, L. S., Moule, C. L., Fraser, H. *et al.* (2005). Natural genetic variants influencing type 1 diabetes in humans and in the NOD mouse. *Novartis Found Symp*, **267**, 57–65; discussion 65–75.

Wordsworth, B. P., Lanchbury, J. S., Sakkas, L. I. *et al.* (1989). HLA-DR4 subtype frequencies in rheumatoid arthritis indicate that DRB1 is the major susceptibility locus within the HLA class II region. *Proc Natl Acad Sci USA*, **86**, 10049–53.

19

Susceptibility to infectious diseases

Andrew J. Walley and Adrian V. S. Hill

Introduction

The study of the host genetic component of infectious disease is potentially one of the most difficult areas of complex human genetic disease analysis for one major reason: the absolute requirement for a host–pathogen interaction to cause disease. Since the pathogen has its own genome, with all of its attendant potential for variability, there has been an ongoing ''arms race'' between man and microbe that has driven each to incorporate changes within their genomes that increase their survival chances should they encounter the other. However, this strong evolutionary pressure on the human genome has also inevitably produced associations between gene variants and disease. The field of genetic susceptibility to infectious disease has been around for over 50 years now and there is a substantial body of evidence for the role of genetics in infectious disease susceptibility.

Historical perspective

Infectious disease is as old as humanity and remains a significant influence on polymorphism in the human genome. Major effects of infectious disease such as epidemics that have drastically reduced populations to a small percentage of individuals, so-called ''bottlenecks,'' such as the Black Death in Europe and the introduction of smallpox and other diseases to the Americas, have a strong selective effect. In addition, the slow continual onslaught of endemic diseases that are still with us today, such as malaria, tuberculosis and, more recently, AIDS, lead to a continual enrichment within the population for resistance alleles, even when these might be harmful and would be quickly lost in the absence of disease.

A recent paper on prion disease illustrates the long-standing effects of the interaction of infectious agents and the human genome (Mead *et al.*, 2003). Prion diseases are transmissible neurodegenerative conditions with a long incubation period that are invariably fatal. Creutzfeld-Jakob disease and kuru are examples of human prion diseases. Heterozygosity for a methionine/valine polymorphism at amino acid 129 of the prion protein confers resistance to prion disease. By comparison to the allele frequencies found in the Fore people of Papua New Guinea, who became endemically infected with kuru after the adoption of cannibalistic funeral rites in the late nineteenth century, the worldwide frequencies of the alleles suggest that balancing selection has been present globally since prehistoric times. The authors propose that acquired prion disease was enough to provide the necessary selective pressure to maintain the variant in the population, either through widespread cannibalistic practices in prehistoric populations or via cross-species transmission from animals.

That infectious disease makes a major contribution to morbidity and mortality across the world is without doubt (see Table 19.1). Although some of

Table 19.1. Leading causes of disability-adjusted life-years (DALYs) and deaths due to infectious disease for the world in 1999

Disease	DALYs (× 1000)	Deaths (× 1000)
Lower respiratory tract infections	96 682	3963
HIV/AIDS	89 819	2673
Diarrheal diseases	72 063	2213
Malaria	44 998	1086
Tuberculosis	33 287	1669
Sexually transmitted diseases (not HIV)	19 747	178

(Adapted from Michaud *et al.*, 2001).

these diseases are now under control in developed countries and eradication has been possible for smallpox and soon polio, the emergence of new diseases such as AIDS, Ebola and West Nile virus mean that studies of genetic susceptibility to infectious disease remain extremely relevant in the fight against disease.

Twin studies and adoption studies

The first formal evidence that genetics is important in susceptibility to infectious disease was gained in the 1930s through work on twins with tuberculosis (Diehl and Von Verschuer, 1936) and twin studies have been carried out subsequently for a number of infectious diseases such as malaria (Sjoberg *et al.*, 1992; Jepson *et al.*, 1994), tuberculosis (Kallmann and Reisner, 1942; Comstock, 1978), leprosy (Chakravarti and Vogel, 1973) and hepatitis B (Lin *et al.*, 1989). These studies all demonstrated a substantial role for genetics in host susceptibility because monozygotic twins were more likely to have concordant disease states than dizygotic twins. Another useful study design for the investigation of genetic effects is the adoptee study, where risk of disease in an adoptee is analyzed according to the characteristics of their biological and adoptive parents. Like the twin study, this can then shed light on how important genetic effects are compared to environmental ones. Such a study demonstrated that host genetic factors were important in susceptibility to fatal infectious diseases in northern Europeans (Sorensen *et al.*, 1988). Although these studies emphasize the importance of genetics, the complexity of the host–pathogen–environment relationship is such that there are currently few useful estimates of the genetic contribution from measuring the increased risk of siblings of susceptible individuals compared to the general population (λs) (Wallace *et al.*, 2003).

Although significant achievements were made in the field before the advent of molecular genetics, it has been the increasing availability of polymorphic genetic markers and the means to readily amplify specific regions of DNA using the polymerase chain reaction (PCR) (Saiki *et al.*, 1985) that have created the conditions for the challenges of complex disease genetics to become practicable. The traditional approach has been to use linkage and segregation analysis in affected families and this has resulted in some notable successes in mapping the chromosomal location of major susceptibility genes in schistosomiasis (Marquet *et al.*, 1996) and leprosy (Siddiqui *et al.*, 2001; Tosh *et al.*, 2002). Analysis of rare syndromes has also shed some light on the pathology of some infectious diseases, where interferon-γ and interleukin-12 deficiencies have been shown to greatly increase the risk of infection by atypical mycobacteria and *Salmonella* species (Jouanguy *et al.*, 1996; Newport *et al.*, 1996; Altare *et al.*, 1998*b*; de Jong *et al.*, 1998). However, the typically episodic nature of acute infectious disease more readily lends itself to case-control study designs, the cases being considered as enriched for the susceptible people within the population at any point in time.

Genetic association studies in case-control groups have been the most productive in terms of the numbers of genes identified as being involved in infectious disease susceptibility because the subjects are easier to recruit and the laboratory work involved may be as little as typing a single

polymorphic marker. Because genetic association is only detectable over much smaller genetic distances than genetic linkage, these studies also have the additional attraction that any marker giving a positive statistically significant result is likely to mean that the disease-causing polymorphism is only a relatively small distance away. However, the main problem is that, unlike family studies, which are typically designed so that they control for non-genetic factors, it is not a simple task to match cases and controls successfully. A precise definition of the phenotype that is being tested for association is essential, typically a set of agreed clinical criteria, but the more clear-cut the phenotype, the better. This then gives the best chance of allocating a subject to the correct group. Statistical analysis techniques can assist retrospectively in correcting for documented potential confounders but nothing can compensate for an initial poor case definition.

Two approaches can be taken in discussing the field of infectious disease susceptibility: either a focus on the diseases that have been investigated, or discussion of the candidate gene variants which have been reported to be associated with disease. Here we give the diseases precedence, with the genes being placed into the context of what is known about the pathological processes they are involved with. However, this chapter cannot provide a comprehensive catalogue of all genes now known to be involved in infectious disease but instead we offer a starting point for anyone wishing to investigate the field further.

The immune system is the body's specialized defense system against pathogens and is an extremely complex set of interacting pathways that are triggered by initial recognition of a pathogen's ''foreign'' molecules followed by a range of responses designed to destroy the pathogen itself, pathogen-infected cells and any molecular species generated by the pathogen. In turn, pathogens have evolved a wide range of mechanisms whereby they attempt to evade immune recognition, e.g. secretion of protein inhibitors to block the immune response, antigenic variation, sequestration inside the host's cells or molecular mimicry of host ''self'' antigens.

The best understood arm of the immune response is the adaptive, or humoral, response. This is extensively described in Chapter 9 and there have been numerous reports of associations of genetic susceptibility to infectious disease with genes in the major histocompatibility complex (MHC) on chromosome 6p21.3. This region contains the human leukocyte antigen (HLA) genes. The function of the HLA proteins is to present small peptides derived from degradation of intracellular proteins, including those from pathogens, on the cell surface where they can be recognised by the host's T-cells. This critical role in triggering the body's response to a pathogen is reflected in the wide range of diseases that variants in HLA alleles are associated with. The HLA Class I proteins (i.e. HLA A, B and C) are expressed on all nucleated cells and present peptides to cytotoxic CD8+ T-cells. HLA Class II proteins (i.e. HLA DP, DQ and DR) are expressed on B-cells, macrophages and antigen-presenting cells and present antigens to helper CD4+ cells. The HLA Class III region encodes a number of complement components (e.g. C2, C4A and C4B) and cytokines (e.g. tumor necrosis factor and the lymphotoxins). The associations between variants in this region and disease are discussed in the relevant sections below.

Recently, the mechanisms of the other arm of the immune system, the innate, or cellular, response have been explored, and one of several groups of molecules recently found to be involved in recognition of pathogen-associated molecular patterns (PAMPs) deserves special mention, the *Toll*-like receptors (TLRs). The ten known human TLRs, their gene location and the molecules and/or pathogens that they recognize are listed in Table 19.2. The identification of this family of receptors stemmed from work carried out in the fruit fly, *Drosophila melanogaster*, where the prototype *Toll* protein was identified (Hashimoto *et al.*, 1988). Initially, it was recognized because of its essential role in development as a determinant of dorsoventral polarity in the fruit fly embryo, but

Table 19.2. Human *Toll*-like receptors, their reported ligands and selected references

Functional receptor	Pathogen-associated molecular patterns (PAMPs)	Selected references
TLR1/TLR2 dimers	*Neisseria meningitidis* soluble factor	(Wyllie *et al.*, 2000)
	Spirochaete glycolipid	(Opitz *et al.*, 2001)
	Staphylococcus phenol-soluble modulin	(Hajjar *et al.*, 2001)
	Borrelia burgdorferi OspA	(Alexopoulou *et al.*, 2002)
	Mycobacterium leprae	(Krutzik *et al.*, 2003)
TLR2	Gram-positive bacteria, e.g. peptidoglycan, lipoteichoic acid	(Yoshimura *et al.*, 1999; Schroder *et al.*, 2003)
	Mycobacteria, e.g. lipoarabinomannan, mycobacterial cell wall	(Brightbill *et al.*, 1999; Means *et al.*, 1999; Underhill *et al.*, 1999*b*)
	Protozoa e.g. *T.cruzii* GPI anchor	(Campos *et al.*, 2001)
	Neisseria porin	(Massari *et al.*, 2002)
	Klebsiella outer membrane protein A	(Jeannin *et al.*, 2000)
	Yeast zymosan	(Underhill *et al.*, 1999*a*)
TLR3	Poly I:C, viral dsRNA	(Alexopoulou *et al.*, 2001)
TLR4	Gram-negative bacteria, e.g. LPS	(Poltorak *et al.*, 1998*a*)
	M. tuberculosis heat sensitive fraction	(Means *et al.*, 1999)
	Respiratory Syncytial Virus fusion protein	(Kurt-Jones *et al.*, 2000)
	Heat shock protein 60	(Ohashi *et al.*, 2000)
	Flavobacteria flavolipin	(Gomi *et al.*, 2002)
TLR5	Bacterial flagellin	(Hayashi *et al.*, 2001)
TLR6/TLR2 Dimers	Gram-positive peptidoglycan	(Ozinsky *et al.*, 2000)
	Mycoplasmal macrophage activating lipopeptide 2 (MALP2)	(Takeuchi *et al.*, 2001)
	Group B streptococcus heat labile soluble factor (GBS-F)	(Henneke *et al.*, 2001)
	Staphylococcus phenol-soluble modulin	(Hajjar *et al.*, 2001)
TLR7	Imiquimod, R-848	(Hemmi *et al.*, 2002; Nagase *et al.*, 2003)
	Loxoribine	(Heil *et al.*, 2003)
	Bropirimine	(Akira *et al.*, 2003)
TLR8	Imiquimod, R-484	(Nagase *et al.*, 2003)
TLR9	CpG bacterial DNA	(Hemmi *et al.*, 2000)
TLR10	None so far	

its crucial role in innate immunity was identified soon afterwards. Flies lacking *Toll* were severely immunocompromised and fungal infections were lethal. This work was then extended to an animal model by the positional cloning of the gene responsible for bacterial lipopolysaccharide (LPS) sensitivity in the *lps* mouse model, a gene identified as a homologue of *Toll*, i.e. *Toll*-like, and now known as *TLR*4 (Poltorak *et al.*, 1998*a*; Poltorak *et al.*, 1998*b*; Hoshino *et al.*, 1999). Other TLRs had been cloned prior to this (Rock *et al.*, 1998) but it was not until the work in the mouse that their function in innate immunity was recognized.

It should be noted however, that it is becoming increasingly clear that the TLRs are not the only PAMP-recognition molecules of importance in disease susceptibility. Recently, as well as NOD family and related intracellular receptors, C-type lectins have been identified as mediators of pathogen recognition as part of innate immunity. The best known example of this is CD209, perhaps better known as DC-SIGN. The gene product was originally isolated on the basis of HIV-1 gp120-binding capability (Curtis *et al.*, 1992) in 1992 but it was not until 2000 that it was recognized as a dendritic-cell-specific receptor for ICAM3 which mediated dendritic-cell and resting T-cell interaction as part of the primary immune response (Geijtenbeek *et al.*, 2000*b*) and consequently, enhanced HIV-1 *trans*-infection of T-cells (Geijtenbeek *et al.*, 2000*a*). Subsequently, DC-SIGN has been shown to mediate dendritic-cell responses to many microorganisms, including: *Aspergillus fumigatus* (Serrano-Gomez *et al.*, 2005), *Streptococcus pneumoniae* (Koppel *et al.*, 2005), measles virus (de Witte *et al.*, 2006) and Ebola virus (Alvarez *et al.*, 2002). Unsurprisingly, variability in this gene has also recently been suggested to contribute to disease susceptibility in tuberculosis (Barreiro *et al.*, 2006) and dengue fever severity (Sakuntabhai *et al.*, 2005).

The range of diseases for which genetic susceptibility has been demonstrated are shown in Table 19.3 with their causative organism, some of the principal candidate genes reported to be associated with them and key references. We focus in this chapter on three diseases that are of major importance for global health and provide good examples of genetic susceptibility to infectious disease. These diseases are malaria, mycobacterial diseases and HIV/AIDS.

Malaria

Malaria provides the classic example of an infectious disease exerting an evolutionary pressure on humans so that protective mutations in genes are maintained in the populations who live in malaria endemic areas. Once the close geographical correlation of the prevalences of malaria and red blood cell disorders (hemoglobinopathies) was noted, Haldane proposed, in the late 1940s, that they might be protective against severe malaria (Haldane, 1949). Only a few years later Allison published the first evidence of a genetic disorder, sickle cell trait, providing a selective advantage against disease (Allison, 1954). The genetic basis of sickle cell trait is a single nucleotide polymorphism (SNP), an adenine to thymine change, which results in a single amino acid change in the hemoglobin molecule — a glutamate at position 6 changes to valine. It is inherited as a monogenic recessive disorder. Homozygotes for the mutation have sickle cell anemia, requiring regular blood transfusions to prevent early death. However, heterozygotes have a 90% reduction in risk of severe malaria and are less likely to die, though they have a similar likelihood of infection as wild-type homozygotes (see Ashley-Koch *et al.*, 2000 for review).

Subsequently, other variants of hemoglobin have been identified with similar effects on malaria susceptibility. Hemoglobin C, a glutamate to lysine change at position 6, is common in West Africa and is less deleterious than HbS and may be expected to displace HbS as the protective variant in time (Modiano *et al.*, 2001). Hemoglobin E, a glutamate to lysine change at position 26, is common in parts of South East Asia (Goueffon and du Saussay, 1969;

Table 19.3. Infectious diseases, reportedly associated genes and selected references. Gene symbols are from the HUGO Gene Nomenclature Committee website (www.gene.ucl.ac.uk/nomenclature). Certain symbols are recent and in order to aid reading the literature their old symbols are as follows: TNF = TNFα, FCGR2 = CD32, FY = DARC, CXCL12 = SDF-1, CX3CR1 = GPR13 or CCRL1, CCL3 = MIP-1α, CCL4 = MIP-1β, CCL5 = RANTES and MBL2 = MBL

Disease	Causative organism	Associated genes	References
PROTOZOA:			
Malaria	*Plasmodium falciparum*	*HBB*	(Allison, 1954; Ashley-Koch *et al.*, 2000)
	Plasmodium vivax		(Modiano *et al.*, 2001)
	Plasmodium ovale		(Hutagalung *et al.*, 1999; Chotivanich *et al.*, 2002)
		G6PD	(Ruwende *et al.*, 1995)
		FY	(Miller *et al.*, 1975)
		GYPC	(Patel *et al.*, 2001)
		SLC4A1	(Jarolim *et al.*, 1991; Mgone *et al.*, 1996)
		HLA	(Hill *et al.*, 1991; Contu *et al.*, 1998)
		TNF	(McGuire *et al.*, 1994; Knight *et al.*, 1999; McGuire *et al.*, 1999)
		NOS2A	(Burgner *et al.*, 1998; Kun *et al.*, 1998)
		CD36	(Aitman *et al.*, 2000; Omi *et al.*, 2003)
		ICAM1	(Fernandez-Reyes *et al.*, 1997)
		IFNAR1	(Aucan *et al.*, 2003)
		IFNGR1	(Koch *et al.*, 2002)
		FCGR2	(Shi *et al.*, 2001; Omi *et al.*, 2002)
		Chr5q31–33	(Rihet *et al.*, 1998*b*; Flori *et al.*, 2003)
Schistosomiasis	*Schistosoma mansoni*	*Chr5q31–33*	(Marquet *et al.*, 1996; Kouriba *et al.*, 2005)
		IL13	
Leishmaniasis	*Leishmania donovani*	*SLC11A1*	(Mohamed *et al.*, 2004)
		IL4	(Mohamed *et al.*, 2003)
		IFNGR1	
	Leishmania chagasi	*TNF*	(Karplus *et al.*, 2002)
BACTERIA:			
Tuberculosis	*Mycobacterium tuberculosis*	*SLC11A1*	(Bellamy *et al.*, 1998; Cervino *et al.*, 2000; Greenwood *et al.*, 2000)
		SFTPA1	(Malik *et al.*, 2006)
		SFTPA2	
		IFNG	(Lio *et al.*, 2002; Lopez-Maderuelo *et al.*, 2003; Rossouw *et al.*, 2003; Tso *et al.*, 2005; Etokebe *et al.*, 2006)
		IL10	(Tso *et al.*, 2005)
		DC-SIGN	(Barreiro *et al.*, 2006)
		VDR	(Bellamy *et al.*, 1999; Wilkinson *et al.*, 2000; Bornman *et al.*, 2004)
		HLA	(Singh *et al.*, 1983; Bothamley *et al.*, 1989; Brahmajothi *et al.*, 1991; Kim *et al.*, 2005; Delgado *et al.*, 2006)
		UBE3A	(Cervino *et al.*, 2002)

Table 19.3 (*cont.*)

Disease	Causative organism	Associated genes	References
		P2RX7	(Li *et al.*, 2002)
		IL8	(Ma *et al.*, 2003)
Leprosy	*Mycobacterium leprae*	*HLA*	(Rea *et al.*, 1976; Serjeantson, 1983; Wang *et al.*, 1999; Shaw *et al.*, 2001)
		VDR	(Roy *et al.*, 1999)
		TNF	(Roy *et al.*, 1997)
		COL3A1	(Kaur *et al.*, 1997)
		TLR2	(Kang *et al.*, 2002)
		CTLA4	(Kaur *et al.*, 1997)
		C4B	(de Messias *et al.*, 1993)
		LAMA2	(Wibawa *et al.*, 2002)
		SLC11A1	(Meisner *et al.*, 2001)
		Chr6q25	(Mira *et al.*, 2003)
		Chr10p13	(Siddiqui *et al.*, 2001; Mira *et al.*, 2003)
		Chr20p12	(Tosh *et al.*, 2002)
		PARK2	(Mira *et al.*, 2004)
		PACRG	
Atypical disseminated mycobacterial disease	*Mycobacterium avium*	*IFNGR1*	(Jouanguy *et al.*, 1996; Newport *et al.*, 1996)
		IFNGR2	(Dorman *et al.*, 1998; Doffinger *et al.*, 2000)
		IL12B	(Altare *et al.*, 1998*b*; Picard *et al.*, 2002)
		IL12RB1	(Altare *et al.*, 1998*a*; Verhagen *et al.*, 2000; Sakai *et al.*, 2001)
		STAT-1	(Dupuis *et al.*, 2001)
Sepsis/septic shock	General septic shock	*TNF*	(Mira *et al.*, 1999; Tang *et al.*, 2000; Barber *et al.*, 2004)
		MBL	(Garred *et al.*, 2003)
		TLR4	(Barber *et al.*, 2004)
	Salmonella sp	*IL12RB1*	(de Jong *et al.*, 1998)
	Neisseria meningitidis	*FCGR2*	(Bredius *et al.*, 1994)
Typhoid	*Salmonella typhi*	*HLA*	(Dunstan *et al.*, 2001; Dharmana *et al.*, 2002)
Stomach Ulcers Gastric cancer	*Helicobacter pylori*	*IL1B/IL1RN* *MPO*	(Hamajima *et al.*, 2001*a*; Rad *et al.*, 2003) (Hamajima *et al.*, 2001*b*)
VIRUSES:			
AIDS	HIV-1	*HLA*	(Jeannet *et al.*, 1989; Kaslow *et al.*, 1996; Tang *et al.*, 1999)
		CCR5	(Dean *et al.*, 1996; Liu *et al.*, 1996; Samson *et al.*, 1996)
		CCR2	(Smith *et al.*, 1997; Anzala *et al.*, 1998; Kostrikis *et al.*, 1998)
		CXCL12	(Mummidi *et al.*, 1998; Winkler *et al.*, 1998; Petersen *et al.*, 2005)
		CX3CR1	(Faure *et al.*, 2000)
		CCL3	(Cocchi *et al.*, 1995)

Table 19.3 (*cont.*)

Disease	Causative organism	Associated genes	References
		CCL4	(Cocchi *et al.*, 1995)
		CCL5	(Cocchi *et al.*, 1995; An *et al.*, 2002)
		IL4	(Nakayama *et al.*, 2000)
		IL10	(Shin *et al.*, 2000)
		MBL2	(Maas *et al.*, 1998)
Hepatitis, cirrhosis and liver cancer	Hepatitis Virus B and C	*HLA*	(Thursz *et al.*, 1995; Yoshizawa *et al.*, 2003; Yoon *et al.*, 2005)
		LDLR	(Hennig *et al.*, 2002)
		IL10/TNF	(Lio *et al.*, 2003)
		VDR	(Bellamy *et al.*, 1999)
Bronchiolitis	Respiratory syncytial virus	*IL8*	(Hull *et al.*, 2001; Heinzmann *et al.*, 2004; Hull *et al.*, 2004; Puthothu *et al.*, 2006)
		TLR4	(Tal *et al.*, 2004)
		IL4	(Hoebee *et al.*, 2003)
		IL4R	
		IL10	(Wilson *et al.*, 2005)
		CCR5	(Hull *et al.*, 2003)
Gastroenteritis	Norwalk virus	*FUT2*	(Lindesmith *et al.*, 2003)
Dengue fever	Dengue virus	*DC-SIGN*	(Sakuntabhai *et al.*, 2005)
PROTEINS:			
Creutzfeld-Jakob disease	Prion protein	*PRNP*	(Goldgaber *et al.*, 1989; Owen *et al.*, 1989; Fink *et al.*, 1991; Mead *et al.*, 2001; Croes *et al.*, 2004)
		PRND	(Croes *et al.*, 2004)
		HLA	(Kuroda *et al.*, 1986)
		APOE	(Amouyel *et al.*, 1994; Van Everbroeck *et al.*, 2001)

Sanguansermsri *et al.*, 1987) and also appears to be associated with reduced disease severity rather than risk of infection (Hutagalung *et al.*, 1999; Chotivanich *et al.*, 2002). Other disorders of globin synthesis, such as the α- and β-thalassemias, are common in malarious regions of the world with studies demonstrating that they too appear to protect against malaria (Flint *et al.*, 1986; Williams *et al.*, 2005). The exact mechanism of protection is, however, unclear and much functional work remains to be done.

Haptoglobin, an acute phase response protein, has been associated with malaria in studies where variant protein isoforms have been typed in study groups (Elagib *et al.*, 1998; Quaye *et al.*, 2000). However, absence of haptoglobin (ahaptoglobinemia) is a common condition in malaria-infected individuals so these phenotypic, rather than genotypic, association studies are complicated by disease status (Rougemont *et al.*, 1980; Boreham *et al.*, 1981; Hill *et al.*, 1987). While this may result in observable correlations between haptoglobin isoforms and disease, the actual association may be to a different gene that is involved in the complex mechanisms that regulate haptoglobin levels. A recent large-scale genotypic investigation in a Gambian case-control study demonstrated no association with this gene (Aucan *et al.*, 2002).

Glucose 6-phosphate dehydrogenase (G6PD) deficiency is the commonest enzyme deficiency in the world affecting an estimated 400 million people (see Ruwende and Hill, 1998 for review). The deficiency renders red blood cells, and some other cell types, unable to regenerate reduced nicotinamide adenine dinucleotide phosphate (NADPH), and this leads to hemolytic anemia and prolonged neonatal jaundice. Hemolytic crises in response to certain foods or medicines are common. The gene is on the X chromosome so males are hemizygous and more likely to have the deficiency than females as it is a recessive disease. However, it remains in the population almost certainly because male hemizygotes and female heterozygotes are resistant to malaria, at least in Africa (Ruwende *et al.*, 1995). Recently, investigation of flanking markers around the *G6PD* and *CD40* ligand genes has provided evidence of recent positive selection by malaria of particular haplotypes (Sabeti *et al.*, 2002; Saunders *et al.*, 2002).

The Duffy blood group is part of the Duffy antigen/chemokine receptor (DARC). A variant of the promoter of this receptor's gene that alters a GATA-1 (a hematopoietic cell-specific transcription factor) binding site inhibits DARC expression on the red blood cell surface (Tournamille *et al.*, 1995) and underlies the Duffy negative phenotype that is very prevalent in sub-Saharan Africa. This then prevents DARC-mediated entry of *Plasmodium vivax* into the red blood cells and so protects from malaria (Miller *et al.*, 1975). However, *P. vivax* very rarely causes fatal disease so malaria might not be the reason for the maintenance of this variant at high frequency in African populations (Hamblin and Di Rienzo, 2000). *CD36* is another gene whose protein is found on a blood cell surface, in this case the platelet as well as on capillary endothelium where it acts as a receptor for *P. falciparum*-infected erythrocytes. However, the relationship of variants in this gene with malaria is complex. One study has shown that *CD36* deficiency is very common in Africans but increases susceptibility to severe malaria (Aitman *et al.*, 2000) while another has demonstrated variants that are associated with protection against severe malaria in Thailand (Omi *et al.*, 2003) suggesting that other factors than malaria are involved in selecting for *CD36* deficiency in Africa.

Different associations have been detected in different ethnic groups affected by malaria, clearly illustrating the effects of ethnicity and local environment. The protein that is the basis of the Gerbich blood group system, Glycophorin C (GYPC), has been demonstrated to be the receptor for the *P. falciparum* erythrocyte binding antigen 140 protein (EBA140). Deletion of exon 3 of the gene has been found at high frequency in coastal areas of Papua New Guinea where malaria is endemic (Patel *et al.*, 2001). It was also demonstrated that lack of GYPC on the erythrocyte surface prevents EBA140 binding and *P. falciparum* invasion via this pathway (Maier *et al.*, 2003). Heterozygosity for a 27bp deletion in the gene for another erythrocyte membrane protein known as band 3 (SLC4A1) has also been associated with protection from severe malaria in Melanesia, even though it causes ovalocytosis, a condition associated with mild hemolytic anemia. It is presumed to be lethal in homozygote form as no homozygotes have ever been detected (Jarolim *et al.*, 1991; Mgone *et al.*, 1996). An association between the intracellular adhesion molecule-1 variant known as $ICAM^{Kilifi}$ and severe malaria was detected in Kenyan children (Fernandez-Reyes *et al.*, 1997), but not in West Africans, with this polymorphism. If this is a real association rather than just a chance finding, it is presumably due to either ethnic differences leading to epistatic effects, to malaria strain variation, or to other complex environmental effects.

The Fulani ethnic group of West Africa are of particular interest as they have been found to be less susceptible to malaria than the neighbouring Mossi and Rimaibé tribes (Modiano *et al.*, 1996). It has been suggested that this may be due to their very high antibody response to malarial antigens and that this may in part be explained by the high frequency of the *-524 T* allele variant of the

Interleukin-4 (IL4) gene promoter, which is believed to enhance production of IL4 (Luoni *et al.*, 2001).

As with most diseases that involve the immune response, genes within the MHC are involved in malaria susceptibility. The Class I variant *HLA-B*53* is associated with resistance to severe disease in West Africa (Hill *et al.*, 1991). In the same study, the Class II variant HLA-*DRB1*1302* was associated with resistance to severe malarial anaemia. *HLA-B*35* frequency has also been associated with malaria endemicity within Sardinia (Contu *et al.*, 1998). The tumor necrosis factor-α (*TNFA*) gene is in the MHC and this has been associated with malaria in a number of studies that attempted to distinguish the *TNFA* effect from those of flanking *HLA* genes. The TNFA protein has a wide range of functions but it was noted in clinical studies that high TNFA levels in the serum correlated with severe malaria (Grau *et al.*, 1989; Shaffer *et al.*, 1991). Subsequently, promoter polymorphisms were demonstrated to be associated with increased risk of disease: some of these may cause increased expression of *TNFA* and one clearly affects the binding of the OCT-1 transcription factor (McGuire *et al.*, 1994; Knight *et al.*, 1999; McGuire *et al.*, 1999).

Various nitric oxide synthase 2A (*NOS2A*) gene polymorphisms have been reported to associate with malaria susceptibility but there is some inconsistency between reports (Burgner *et al.*, 1998; Kun *et al.*, 1998; 2001). One variant was found to increase production of the enzyme and hence nitric oxide, and appeared protective against malaria in two groups from Tanzania and Kenya (Hobbs *et al.*, 2002).

More recently, various receptor gene variants have been investigated as candidate genes for malaria susceptibility. The receptors for two of the interferons have both been associated with malaria in Africans. The interferon-α receptor-1 (*IFNAR1*) gene has been associated with protection against severe malaria and cerebral malaria in particular (Aucan *et al.*, 2003). Polymorphisms in the promoter region of the interferon-γ receptor-1 gene have also been associated with cerebral malaria (Koch *et al.*, 2002). An amino acid change in the Fcγ receptor IIA (*CD32*) gene has been reported to be associated with protection against high levels of parasitemia in the first year of life in a Western Kenyan study (Shi *et al.*, 2001) and to susceptibility to malaria in Thailand (Omi *et al.*, 2002). We have recently observed an association with severe malaria disease in a Gambian case-control study (Cooke *et al.*, 2003). Another study on parasitemia levels moved from linkage data in families (Rihet *et al.*, 1998*a*) to association with specific markers in the chromosome 5q31–33 cytokine cluster (Rihet *et al.*, 1998*b*), though no causative gene has been identified as yet.

Extrapolation from mouse models of malaria to human disease is complicated by a large genetic distance between the human and murine malaria parasites and pathophysiological differences, particularly with respect to cerebral malaria which is not really observed in mice. However, the mouse is very well understood in terms of its genetics, and attempts have been made to try and locate regions that are linked to malaria susceptibility. Susceptibility to infection with *Plasmodium chabaudi* varies from mouse strain to mouse strain and using standard informative backcross and F2 group comparisons, four *P. chabaudi* resistance (Char) loci have been mapped on mouse chromosomes 9 (*Char1*), 8 (*Char2*), 17 (*Char3*) and 3 (*Char4*) (see Fortin *et al.*, 2002 for review). However, as yet, none of these mouse linkages has been taken forward to identify a human susceptibility gene.

Mycobacterial disease: tuberculosis

Tuberculosis is one of the leading causes of bacterial mortality worldwide (Table 19.1). Infection occurs following inhalation of *Mycobacterium tuberculosis*. The majority of bacteria are probably killed immediately with a few surviving inside macrophages. In 90–95% of cases, the disease does not progress any further;

however, in the remaining 5–10% of people they multiply and the disease tuberculosis develops. Accidental injection of 249 babies in Lubeck, Germany, in 1926 with virulent *M. tuberculosis* instead of the vaccine strain Bacille Calmette-Guerin (BCG) resulted in 173 surviving, indicating effective but variable innate immunity to tuberculosis in these babies (Diehl and Von Verschuer, 1936). A number of twin studies have demonstrated a significant genetic component to tuberculosis susceptibility (Kallmann and Reisner, 1942; Comstock, 1978; Jepson *et al.*, 2001).

A mouse model provided some insight into genetic susceptibility to mycobacterial disease when a single genetic factor which rendered mice resistant to *M. bovis* (BCG) was identified (Gros *et al.*, 1981). Linkage studies and positional cloning mapped and identified the gene on mouse chromosome 1 (Vidal *et al.*, 1993). This gene was initially designated the natural resistance-associated macrophage protein gene (*NRAMP1*) in humans but is now known as solute carrier family 11 member A1 (SLC11A1) and is located on human chromosome 2q35. SLC11A1 has been associated with pulmonary tuberculosis in a number of different populations but not in others, providing evidence that it probably plays a variable role in human susceptibility (Bellamy *et al.*, 1998; Cervino *et al.*, 2000; Greenwood *et al.*, 2000; Malik *et al.*, 2005). The exact role of SLC11A1 in the pathogenesis of tuberculosis is uncertain, although it is known that the main function of the protein is to transport divalent cations across the phagolysosomal membrane. There is also some evidence for a role for SLC11A1 in visceral leishmaniasis (Mohamed *et al.*, 2004).

Another insight into susceptibility to tuberculosis arose from the observation that vitamin D was reported to be clinically efficacious in the treatment of cutaneous tuberculosis (Dowling *et al.*, 1946). Vitamin D deficiency had been linked to tuberculosis from epidemiological evidence as well, with serum 1,25 dihydroxyvitamin D_3 [$1,25(OH)_2D_3$] levels being lower in untreated tuberculosis patients in the UK (Davies *et al.*, 1985) and associated with disease severity in Indonesian patients (Grange *et al.*, 1985). Since $1,25(OH)_2D_3$ exerts its effects via the vitamin D receptor (VDR), variants of this gene might be expected to be associated with susceptibility and this was initially observed in Gambians (Bellamy *et al.*, 1999) and further associated in a study looking at case-control and family cohorts from The Gambia, Guinea and Guinea-Bissau (Bornman *et al.*, 2004). The association was also observed in an ethnically distinct group of Gujarati tuberculosis patients in the UK (Wilkinson *et al.*, 2000).

Studies of adaptive and innate immunity genes have shown associations to both arms of the immune response. Several studies have identified *HLA* allelic associations with tuberculosis, specifically *HLA-DR2* (Singh *et al.*, 1983; Bothamley *et al.*, 1989; Brahmajothi *et al.*, 1991), though this is not present in all populations studied (Cox *et al.*, 1988). *HLA-DQB1*0503* appears to associated with tuberculosis in Cambodia. (Goldfeld *et al.*, 1998; Delgado *et al.*, 2006). *HLA-DRB1*0803*, and the linked *HLA-DQB1*0601*, were associated with disease progression, the development of drug resistance, severe disease, and recurrent disease in Korean subjects (Kim *et al.*, 2005).

TLR2 has been associated with susceptibility in a mouse model where mice with the *TLR2* gene knocked out, but not mice with knockouts for TLR4 or CD14, were more susceptible to high-dose aerosol challenge than control mice (Reiling *et al.*, 2002). Another PAMP-receptor, DC-SIGN, has very recently been reported to contain a promoter polymorphism that was associated with tuberculosis susceptibility in South Africans (Barreiro *et al.*, 2006).

A specific mycobactericidal activity has also been elucidated whereby ATP-mediated ligation of $P2X_7$ purinergic receptors on macrophages triggers them to rapidly kill mycobacteria. A polymorphism in the promoter of the gene for the $P2X_7$ receptor was found to be associated with protection against tuberculosis, though the

exact mechanism remains to be studied (Li *et al.*, 2002). Similarly, one of the major defences against bacterial pathogens in the lung, Surfactant Protein A (SFTPA1 and 2), has been reported to carry polymorphisms which are associated with susceptibility (Malik *et al.*, 2006).

A genome-wide linkage study using a total of 173 sib pairs from The Gambia and South Africa identified two regions of suggestive linkage on chromosomes 15 and X (Bellamy *et al.*, 2000). Fine-mapping of the 10cM region of linkage on chromosome 15 narrowed this linkage down to implicate a variant of the *UBE3A* gene but the role of this ubiquitin ligase in tuberculosis is currently unknown (Cervino *et al.*, 2002).

Possibly the most reproducible evidence of association with tuberculosis has emerged recently from studies of the interferon-gamma gene. This cytokine plays a key role in defence against mycobacteria in many animal models. Five studies have now reported evidence that the A/T polymorphism at the +874 position of the *IFNG* gene is associated with tuberculous disease. A small Sicilian study (of only 45 patients) found the +874TT genotype associated with protection from chronic tuberculosis (Lio *et al.*, 2002). A larger study of Spanish cases found a significant susceptibility effect for the +874AA genotype. Also peripheral blood mononuclear cells from individuals with the +874AA genotype produced less interferon-γ when stimulated by a Purified Protein Derivative (PPD) antigen (Lopez-Maderuelo *et al.*, 2003). In the Cape Colored population of South Africa, the +874AA genotype was associated with tuberculosis in a case-control study and the +874A allele was over-transmitted to affected offspring in families (Rossouw *et al.*, 2003). In Hong Kong Chinese, the +874AA genotype was associated with susceptibility (Tso *et al.*, 2005) and a study of Croatians suggested that the association was to disease severity and not susceptibility per se (Etokebe *et al.*, 2006). The +874 polymorphism alters a binding site for the transcription factor NFκB, and is therefore likely to influence inflammatory responses (Pravica *et al.*, 2000).

Mycobacterial disease: leprosy

Unlike tuberculosis, leprosy is not a disease with a significant mortality rate, but with sometimes devastating effects on nerves and skin, it can have severe consequences on quality of life and significant social and economic effects on a community. Leprosy consists of two forms: tuberculoid and lepromatous. Tuberculoid is characterized by strong cell-mediated immunity, a Th1 CD4+ cytokine profile (interferon-γ, interleukin-12) and very few bacteria in skin lesions. Lepromatous on the other hand is characterized by a lack of cell-mediated immunity, Th2 CD4+ cytokines (interleukins-4 and -5), large numbers of bacteria and progressive disseminated disease. *M. leprae* typically infects macrophages and Schwann cells and like tuberculosis, the majority of people infected probably do not develop any symptoms.

Twin studies and segregation analyses have demonstrated that there is a significant genetic component in leprosy (Chakravarti and Vogel, 1973; Haile *et al.*, 1985; Abel *et al.*, 1988; Wagener *et al.*, 1988). In common with tuberculosis, allelic variants of *SLC11A1*, *VDR* and *HLA-DR* have been associated with leprosy or leprosy type. The MHC contains a substantial number of genes that have been linked to leprosy, though the data should be interpreted with some caution in light of the strong linkage disequilibrium that exists across the region, though it is clear that the evidence linking this region with leprosy susceptibility is strong.

Initial studies demonstrated the non-random segregation of HLA haplotypes amongst children with leprosy (de Vries *et al.*, 1976; Fine *et al.*, 1979). Subsequently the MHC region has been extensively investigated. Several studies have found associations between HLA Class I gene frequencies and either leprosy per se or to one or the other type of leprosy but these have yet to be successfully replicated (Rea *et al.*, 1976; Greiner *et al.*, 1978; Serjeantson, 1983; Rani *et al.*, 1992). An HLA-B46, MICA-5A5 (MHC class I related chain-related

gene A) haplotype has also been associated with protection in a small Chinese study (Wang *et al.*, 1999). Class II genes particularly *HLA-DR2* and *HLA DQw1*, have been associated with leprosy (Miyanaga *et al.*, 1981; Schauf *et al.*, 1985; Mehra *et al.*, 1995; Shaw *et al.*, 2001) as well as the *TAP2* gene (transporter 2, ATP-binding cassette, subfamily B) (Rajalingam *et al.*, 1997) in tuberculoid leprosy patients from North India. A range of genes in the class III region have also been associated with leprosy in a number of studies, particularly TNFa (Roy *et al.*, 1997) but also complement component 4B (*C4B*) (de Messias *et al.*, 1993) and heat shock 70 kD protein 1A (*HSPA1A*) (Rajalingam *et al.*, 2000).

In addition, at least three non-MHC genes have been identified as associated with leprosy: the collagen gene *COL3A1*, the cytotoxic T-lymphocyte activator-4 (*CTLA4*) (Kaur *et al.*, 1997) and laminin-α2 (*LAMA2*) (Wibawa *et al.*, 2002) but these remain to be confirmed. An association to the TLR2 SNP R677W was observed to lepromatous, and not tuberculoid, leprosy in a small sample of Korean patients (Kang *et al.*, 2002). Although the genetic association has not been replicated yet, it has been demonstrated that the variant allele is unable to mediate the cellular signaling response that would normally result in NF-κB activation (Krutzik *et al.*, 2003).

Two successful genome screens have been carried out on affected sib pairs with leprosy so that this is arguably the infectious disease that has been most successfully tackled by family-based linkage analysis. The first initially identified a linkage to chromosome 10p13 (Siddiqui *et al.*, 2001) in Indian subjects and then by typing additional markers a second locus on chromosome 20p12 was identified (Tosh *et al.*, 2002). Intriguingly, this was in a similar position to peaks of linkage identified in genome screens for atopic dermatitis and psoriasis, two disorders also characterized by immune system mediated damage of the skin (Trembath *et al.*, 1997; Cookson *et al.*, 2001). The second genome screen in Vietnamese subjects confirmed the chromosome 10p13 linkage and reported a novel linkage to the chromosomal region 6q25 (Mira *et al.*, 2003). This region of linkage was then narrowed down using association studies to a shared regulatory region between the PARK2 (Parkinson disease 2) and PACRG (PARK2 co-regulated) genes (Mira *et al.*, 2004). Also, the gene and causative amino acid changes underlying the Indian linkage peak on chromosome 10p13 have been identified (Tosh *et al.*, unpublished).

Mycobacterial disease: rare atypical infections

Although genetic susceptibility to infectious disease is complex, as in many other complex genetic traits, the study of rare Mendelian disorders that give rise to similar phenotypes can be useful in understanding the underlying pathogenic mechanisms of a disease. Sporadic cases of disseminated mycobacterial infections occurring in more than one family member have been reported a number of times in the literature (Levin *et al.*, 1995; Casanova *et al.*, 1996). These infections are atypical because they involve typically poorly pathogenic mycobacterium such as *M. avium*, *M. cheloniae* and *M. fortuitum*.

In some families the molecular defect has now been identified. These include mutations in the genes for the interferon-γ receptors 1 and 2, the interleukin-12 p40 subunit and the interleukin-12 receptor β1 subunit (see Table 19.3 for references). Two clinical subtypes have been defined and are related to the type of deficiency caused. Complete loss of the *IFNGR1* or *2* gene products results in an overwhelming infection and impaired granuloma formation, while partial deficiency of any of the four genes or complete deficiency of the *IL12p40* or *IL12RB1* genes leads to a milder curable infection with mature granuloma formation.

Human immunodeficiency virus and acquired immune deficiency syndrome (HIV and AIDS)

HIV-1 infection is characterized by a clinically asymptomatic phase that eventually leads to immune deficiency, opportunistic infections, neurological symptoms and malignancies which constitute AIDS. The clinical course of HIV-1 infection is highly variable, with untreated individuals developing AIDS from 2–3 years after seroconversion or remaining asymptomatic for over 15 years (Hendriks *et al.*, 1993). The first gene to be associated with resistance to HIV infection was isolated when individuals who had been frequently exposed to HIV-1 but were uninfected, were identified. The gene was the chemokine receptor *CCR5* and the resistance was conferred by being homozygous for a 32 bp deletion (*CCR5Δ32*) within the coding region (Dean *et al.*, 1996; Liu *et al.*, 1996; Samson *et al.*, 1996). This deletion leads to a premature frame shift and a non-functional protein that is not expressed on the cell surface. The discovery of the involvement of CCR5 led to a detailed elucidation of the mechanism of entry of HIV-1 via the envelope glycoprotein gp120 binding to CD4 and then a chemokine receptor, eventually leading to the fusion of the viral and cellular membranes. It is notable that the *CCR5Δ32* polymorphism is not present in Africans, presumably due to the absence of the selection pressure of whatever is preserving the polymorphism in Western populations (Ramaley *et al.*, 2002). Since HIV-1 has been present in the West only in recent years it cannot be responsible for the high frequency of this allele and other infectious agents such as *Yersinia pestis* have been suggested.

The *CCR5Δ32* mutation is also associated with a reduction in the rate of progression to AIDS in heterozygotes (Huang *et al.*, 1996), presumably due to the reduction in CCR5+ cells available to the virus on initial infection. Conversely, a common haplotype in the CCR5 promoter, consisting of a combination of 10 SNPs and known as P1, is associated with more rapid progression of the disease (Martin *et al.*, 1998). It is likely that this is because one or more of the SNPs, or one or more unidentified SNPs in linkage disequilibrium with them, are involved in regulating the transcription of CCR5 and hence the expression levels on the cell surface. Such a SNP has been recently proposed; the promoter SNP -2459 A/G has been demonstrated to affect CCR5 expression on the cell surface and hence HIV-1 disease progression (Salkowitz *et al.*, 2003).

Further work on chemokine receptors identified additional chemokine co-receptors, such as CXCR4, which is particularly important as the co-receptor for syncytium-inducing (SI) HIV-1 variants, whereas CCR5 is the co-receptor for non-syncytium-inducing (NSI) variants. These are of clinical relevance as the initial infection is generally established by NSI variants, which are macrophage-tropic due to the requirement for CCR5. SI variants then subsequently emerge in around half of all HIV patients and these can target memory and naïve CD4+ T cells leading to rapid CD4+ T cell loss. No genetic association with this receptor has yet been confirmed though. A CCR2 SNP, corresponding to a valine to isoleucine change at postion 641 in the protein, has also been associated with delayed progression to AIDS in several cohorts but the mechanism is unclear as CCR2 is rarely used as a co-receptor (Smith *et al.*, 1997; Anzala *et al.*, 1998; Kostrikis *et al.*, 1998). A CX3CR1 SNP has been reported to be associated with enhanced progression to AIDS but this effect has not so far been replicated (Faure *et al.*, 2000).

Given that chemokine receptors are essential for successful HIV-1 infection, it is not surprising that the chemokine receptor ligands have also been identified as associated with HIV susceptibility, though it should be noted that replication of these studies has been less successful. CCR5 ligands, such as CCL5 (RANTES, regulated upon activation, Normally T-cell expressed, and presumably Secreted), CCL3 (macrophage inflammatory protein-1 α), and CCL4 (Macrophage inflammatory protein-1 β) have been shown to

block HIV-1 infection in vitro (Cocchi *et al.*, 1995). A 2-SNP CCL5 haplotype in the promoter was associated with reduced progression of disease and both of the SNPs displayed increased promoter activity (An *et al.*, 2002). A protective effect was also reported for a SNP in the *CXCL12* gene (stromal cell-derived factor 1), which is a ligand for the CXCR4 receptor. However, different studies of the *CXCL12* variant have produced some striking discrepancies (Mummidi *et al.*, 1998; Winkler *et al.*, 1998; Petersen *et al.*, 2005).

HIV infection and disease progression has also been demonstrated to be associated with variants of *HLA* genes. The class I alleles *B35* and *Cw4* have been associated with enhanced progression (Jeannet *et al.*, 1989; Tomiyama *et al.*, 1997; Carrington *et al.*, 1999) and *B14, B27, B57* and *C8* have all been associated with non-progression (Kaslow *et al.*, 1996; Hendel *et al.*, 1999; Keet *et al.*, 1999; Migueles *et al.*, 2000). There is also evidence that a heterozygous advantage operates for HLA Class I genes (Keet *et al.*, 1999; Tang *et al.*, 1999). The theory is that heterozygosity of HLA class I loci confers an advantage because it means that the individual can present the maximum possible repertoire of antigenic epitopes. Homozygosity at one or more *HLA* class I loci was strongly associated with rapid progression in both Caucasians and African-Americans, with each locus contributing independently. Replication of this result for *HLA-A* and *-B* in Dutch men and Rwandan women supported the association with homozygosity (Tang *et al.*, 1999).

Other genes reported to be associated with HIV/AIDS include interleukins -4 and -10 (*IL4, IL10*) and mannose-binding lectin (*MBL*). A promoter variant of interleukin-10, which inhibits secretion of a wide range of inflammatory cytokines and chemokines and reduces infection of macrophages in vitro, has been associated with enhanced progression to AIDS as have interleukin-4 promoter variants (Nakayama *et al.*, 2000; Shin *et al.*, 2000; Smolnikova and Konekov, 2002). A variant of the mannose-binding lectin gene has been reported to be associated with slower progression to AIDS (Maas *et al.*, 1998) but a recent report found no association between MBL levels and HIV disease progression (Malik *et al.*, 2003).

Conclusion

It is clear that there are common genes which are associated with a wide range of infectious diseases as well as disease-specific susceptibility alleles, notably those of the MHC. MHC molecules are most frequently implicated and these are central to the immune response to pathogens. However, each pathogen utilizes different mechanisms to evade immune surveillance and reproduce within the host. The cell types that they infect and the range of mechanisms of antigenic variation which they utilize require a highly sophisticated host response with a broad recognition repertoire and the HLA and T-cell receptor (TCR) molecules provide this. More general systems that recognize broad ranges of molecular patterns such as the TLRs and NOD receptors provide an initial rapid response to pathogenic organisms and may provide critical information on the type and magnitude of the immune response to be generated.

More ambitious genome-wide association studies are currently being initiated on a larger scale than those that were undertaken by genetic linkage. These scans will have the power to detect alleles of smaller effect than the few major loci mappable by linkage studies. With current genotyping technologies, such as the Sequenom mass spectrometry system, and several array-based platforms, high-throughput genotyping of SNPs in large numbers of samples is now possible and the first high-density SNP mapping of the whole genome for association to infectious disease is now in progress for malaria and tuberculosis. Such an approach is being widely adopted for many complex diseases but should be particularly informative in infectious diseases, which are generally highly polygenic and we can look forward to a wealth of new information on susceptibility genes.

REFERENCES

Abel, L. and Demenais, F. (1988). Detection of major genes for susceptibility to leprosy and its subtypes in a Caribbean island: Desirade island. *Am J Hum Genet*, **42**, 256–66.

Aitman, T. J., Cooper, L. D., Norsworthy, P. J. *et al.* (2000). Malaria susceptibility and CD36 mutation. *Nature*, **405**, 1015–16.

Akira, S. and Hemmi, H. (2003). Recognition of pathogen-associated molecular patterns by TLR family. *Immunol Lett*, **85**, 85–95.

Alexopoulou, L., Holt, A. C., Medzhitov, R. and Flavell, R. A. (2001). Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. *Nature*, **413**, 732–8.

Alexopoulou, L., Thomas, V., Schnare, M. *et al.* (2002). Hyporesponsiveness to vaccination with *Borrelia burgdorferi OspA* in humans and in TLR1- and TLR2-deficient mice. *Nat Med*, **8**, 878–84.

Allison, A. C. (1954). Protection afforded by sickle-cell trait against subtertian malarial infection. *Br Med J*, **4857**, 290–4.

Altare, F., Durandy, A., Lammas, D. *et al.* (1998*a*). Impairment of mycobacterial immunity in human interleukin-12 receptor deficiency. *Science*, **280**, 1432–5.

Altare, F., Lammas, D., Revy, P. *et al.* (1998*b*). Inherited interleukin 12 deficiency in a child with bacille Calmette-Guerin and *Salmonella enteritidis* disseminated infection. *J Clin Invest*, **102**, 2035–40.

Alvarez, C. P., Lasala, F., Carrillo, J. *et al.* (2002). C-type lectins DC-SIGN and L-SIGN mediate cellular entry by Ebola virus in *cis* and in *trans*. *J Virol*, **76**, 6841–4.

Amouyel, P., Vidal, O., Launay, J. M. and Laplanche, J. L. (1994). The apolipoprotein E alleles as major susceptibility factors for Creutzfeldt-Jakob disease. The French Research Group on Epidemiology of Human Spongiform Encephalopathies. *Lancet*, **344**, 1315–18.

An, P., Nelson, G. W., Wang, L. *et al.* (2002). Modulating influence on HIV/AIDS by interacting RANTES gene variants. *Proc Natl Acad Sci USA*, **99**, 10002–7.

Anzala, A. O., Ball, T. B., Rostron, T. *et al.* (1998). CCR2–64I allele and genotype association with delayed AIDS progression in African women. University of Nairobi Collaboration for HIV Research. *Lancet*, **351**, 1632–3.

Ashley-Koch, A., Yang, Q. and Olney, R. S. (2000). Sickle hemoglobin (HbS) allele and sickle cell disease: a HuGE review. *Am J Epidemiol*, **151**, 839–45.

Aucan, C., Walley, A. J., Greenwood, B. M. and Hill, A. V. (2002). Haptoglobin genotypes are not associated with resistance to severe malaria in The Gambia. *Trans R Soc Trop Med Hyg*, **96**, 327–8.

Aucan, C., Walley, A. J., Hennig, B. J. *et al.* (2003). Interferon-alpha receptor-1 (*IFNAR1*) variants are associated with protection against cerebral malaria in The Gambia. *Genes Immun*, **4**, 275–82.

Barber, R. C., Aragaki, C. C., Rivera-Chavez, F. A. *et al.* (2004). TLR4 and TNF-alpha polymorphisms are associated with an increased risk for severe sepsis following burn injury. *J Med Genet*, **41**, 808–13.

Barreiro, L. B., Neyrolles, O., Babb, C. L. *et al.* (2006). Promoter variation in the DC-SIGN-encoding gene CD209 is associated with tuberculosis. *PLoS Med*, **3**, e20.

Bellamy, R., Beyers, N., McAdam, K. P. *et al.* (2000). Genetic susceptibility to tuberculosis in Africans: a genome-wide scan. *Proc Natl Acad Sci USA*, **97**, 8005–9.

Bellamy, R., Ruwende, C., Corrah, T. *et al.* (1999). Tuberculosis and chronic hepatitis B virus infection in Africans and variation in the vitamin D receptor gene. *J Infect Dis*, **179**, 721–4.

Bellamy, R., Ruwende, C., Corrah, T. *et al.* (1998). Variations in the *NRAMP1* gene and susceptibility to tuberculosis in West Africans. *N Engl J Med*, **338**, 640–4.

Boreham, P. F., Lenahan, J. K., Port, G. R. and McGregor, I. A. (1981). Haptoglobin polymorphism and its relationship to malaria infections in The Gambia. *Trans R Soc Trop Med Hyg*, **75**, 193–200.

Bornman, L., Campbell, S. J., Fielding, K. *et al.* (2004). Vitamin D receptor polymorphisms and susceptibility to tuberculosis in West Africa: a case-control and family study. *J Infect Dis*, **190**, 1631–41.

Bothamley, G. H., Beck, J. S., Schreuder, G. M. *et al.* (1989). Association of tuberculosis and M. tuberculosis-specific antibody levels with HLA. *J Infect Dis*, **159**, 549–55.

Brahmajothi, V., Pitchappan, R. M., Kakkanaiah, V. N. *et al.* (1991). Association of pulmonary tuberculosis and HLA in south India. *Tubercle*, **72**, 123–32.

Bredius, R. G., Derkx, B. H., Fijen, C. A. *et al.* (1994). Fc gamma receptor IIa (CD32) polymorphism in fulminant meningococcal septic shock in children. *J Infect Dis*, **170**, 848–53.

Brightbill, H. D., Libraty, D. H., Krutzik, S. R. *et al.* (1999). Host defense mechanisms triggered by microbial lipoproteins through Toll-like receptors. *Science*, **285**, 732–6.

Burgner, D., Xu, W., Rockett, K. *et al.* (1998). Inducible nitric oxide synthase polymorphism and fatal cerebral malaria. *Lancet*, **352**, 1193–4.

Campos, M.A., Almeida, I.C., Takeuchi, O. *et al.* (2001). Activation of Toll-like receptor-2 by glycosylphosphatidylinositol anchors from a protozoan parasite. *J Immunol*, **167**, 416–23.

Carrington, M., Nelson, G.W., Martin, M.P. *et al.* (1999). HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. *Science*, **283**, 1748–52.

Casanova, J.L., Blanche, S., Emile, J.F. *et al.* (1996). Idiopathic disseminated bacillus Calmette-Guerin infection: a French national retrospective study. *Pediatrics*, **98**, 774–8.

Cervino, A.C., Lakiss, S., Sow, O. *et al.* (2002). Fine mapping of a putative tuberculosis-susceptibility locus on chromosome 15q11–13 in African families. *Hum Mol Genet*, **11**, 1599–603.

Cervino, A.C., Lakiss, S., Sow, O. and Hill, A.V. (2000). Allelic association between the *NRAMP1* gene and susceptibility to tuberculosis in Guinea-Conakry. *Ann Hum Genet*, **64**, 507–12.

Chakravarti, M.R. and Vogel, F. (1973). A twin study in leprosy. *Top Hum Genet*, **1**, 1–123.

Chotivanich, K., Udomsangpetch, R., Pattanapanyasat, K. *et al.* (2002). Hemoglobin E: a balanced polymorphism protective against high parasitemias and thus severe *P falciparum malaria*. *Blood*, **100**, 1172–6.

Cocchi, F., DeVico, A.L., Garzino-Demo, A. *et al.* (1995). Identification of RANTES, MIP-1 alpha, and MIP-1 beta as the major HIV-suppressive factors produced by CD8+ T cells. *Science*, **270**, 1811–15.

Comstock, G.W. (1978). Tuberculosis in twins: a re-analysis of the Prophit survey. *Am Rev Resp Dis*, **117**, 621–4.

Contu, L., Carcassi, C., Orru, S. *et al.* (1998). HLA-B35 frequency variations correlate with malaria infection in Sardinia. *Tissue Antigens*, **52**, 452–61.

Cooke, G.S., Aucan, C., Walley, A.J. *et al.* (2003). Association of Fcgamma receptor IIa (CD32) polymorphism with severe malaria in West Africa. *Am J Trop Med Hyg*, **69**, 565–8.

Cookson, W.O., Ubhi, B., Lawrence, R. *et al.* (2001). Genetic linkage of childhood atopic dermatitis to psoriasis susceptibility loci. *Nat Genet*, **27**, 372–3.

Cox, R.A., Downs, M., Neimes, R.E. *et al.* (1988). Immunogenetic analysis of human tuberculosis. *J Infect Dis*, **158**, 1302–8.

Croes, E.A., Alizadeh, B.Z., Bertoli-Avella, A.M. *et al.* (2004). Polymorphisms in the prion protein gene and in the doppel gene increase susceptibility for Creutzfeldt-Jakob disease. *Eur J Hum Genet*, **12**, 389–94.

Curtis, B.M., Scharnowske, S. and Watson, A.J. (1992). Sequence and expression of a membrane-associated C-type lectin that exhibits CD4-independent binding of human immunodeficiency virus envelope glycoprotein gp120. *Proc Natl Acad Sci USA*, **89**, 8356–60.

Davies, P.D., Brown, R.C. and Woodhead, J.S. (1985). Serum concentrations of vitamin D metabolites in untreated tuberculosis. *Thorax*, **40**, 187–90.

de Jong, R., Altare, F., Haagen, I.A. *et al.* (1998). Severe mycobacterial and Salmonella infections in interleukin-12 receptor-deficient patients. *Science*, **280**, 1435–8.

de Messias, I.J., Santamaria, J., Brenden, M., Reis, A. and Mauff, G. (1993). Association of C4B deficiency (C4B*Q0) with erythema nodosum in leprosy. *Clin Exp Immunol*, **92**, 284–7.

de Vries, R.R., Fat, R.F., Nijenhuis, L.E. and van Rood, J.J. (1976). HLA-linked genetic control of host response to *Mycobacterium leprae*. *Lancet*, **2**, 1328–30.

de Witte, L., Abt, M., Schneider-Schaulies, S., van Kooyk, Y. and Geijtenbeek, T.B. (2006). Measles virus targets DC-SIGN to enhance dendritic cell infection. *J Virol*, **80**, 3477–86.

Dean, M., Carrington, M., Winkler, C. *et al.* (1996). Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. *Science*, **273**, 1856–62.

Delgado, J.C., Baena, A., Thim, S. and Goldfeld, A.E. (2006). Aspartic acid homozygosity at codon 57 of HLA-DQ beta is associated with susceptibility to pulmonary tuberculosis in Cambodia. *J Immunol*, **176**, 1090–7.

Dharmana, E., Joosten, I., Tijssen, H.J. *et al.* (2002). HLA-DRB1*12 is associated with protection against complicated typhoid fever, independent of tumour necrosis factor alpha. *Eur J Immunogenet*, **29**, 297–300.

Diehl, K. and Von Verschuer, O. (1936). Der Erbeinfluss bei den Tuberkulose. *Beitr Klin Kunsch*, **92**, 275.

Doffinger, R., Jouanguy, E., Dupuis, S. *et al.* (2000). Partial interferon-gamma receptor signaling chain deficiency in a patient with bacille Calmette-Guerin and *Mycobacterium abscessus* infection. *J Infect Dis*, **181**, 379–84.

Dorman, S. E. and Holland, S. M. (1998). Mutation in the signal-transducing chain of the interferon-gamma receptor and susceptibility to mycobacterial infection. *J Clin Invest*, **101**, 2364–9.

Dowling, G. B. and Prosser-Thomas, E. W. (1946). Treatment of *lupus vulgaris* with calciferol. *Lancet*, **i**, 919–22.

Dunstan, S. J., Stephens, H. A., Blackwell, J. M. *et al.* (2001). Genes of the class II and class III major histocompatibility complex are associated with typhoid fever in Vietnam. *J Infect Dis*, **183**, 261–8.

Dupuis, S., Dargemont, C., Fieschi, C. *et al.* (2001). Impairment of mycobacterial but not viral immunity by a germline human STAT1 mutation. *Science*, **293**, 300–3.

Elagib, A. A., Kider, A. O., Akerstrom, B. and Elbashir, M. I. (1998). Association of the haptoglobin phenotype (1–1) with falciparum malaria in Sudan. *Trans R Soc Trop Med Hyg*, **92**, 309–11.

Etokebe, G. E., Bulat-Kardum, L., Johansen, M. S. *et al.* (2006). Interferon-gamma gene (T874A and G2109A) polymorphisms are associated with microscopy-positive tuberculosis. *Scand J Immunol*, **63**, 136–41.

Faure, S., Meyer, L., Costagliola, D. *et al.* (2000). Rapid progression to AIDS in HIV+ individuals with a structural variant of the chemokine receptor CX3CR1. *Science*, **287**, 2274–7.

Fernandez-Reyes, D., Craig, A. G., Kyes, S. A. *et al.* (1997). A high frequency African coding polymorphism in the N-terminal domain of ICAM-1 predisposing to cerebral malaria in Kenya. *Hum Mol Genet*, **6**, 1357–60.

Fine, P. E., Wolf, E., Pritchard, J. *et al.* (1979). HLA-linked genes and leprosy: a family study in Karigiri, South India. *J Infect Dis*, **140**, 152–61.

Fink, J. K., Warren, J. T., Jr., Drury, I., Murman, D. and Peacock, M. L. (1991). Allele-specific sequencing confirms novel prion gene polymorphism in Creutzfeldt-Jakob disease. *Neurology*, **41**, 1647–50.

Flint, J., Hill, A. V., Bowden, D. K. *et al.* (1986). High frequencies of alpha-thalassaemia are the result of natural selection by malaria. *Nature*, **321**, 744–50.

Flori, L., Kumulungui, B., Aucan, C. *et al.* (2003). Linkage and association between *Plasmodium falciparum* blood infection levels and chromosome 5q31-q33. *Genes Immun*, **4**, 265–8.

Fortin, A., Stevenson, M. M. and Gros, P. (2002). Complex genetic control of susceptibility to malaria in mice. *Genes Immun*, **3**, 177–86.

Garred, P., Strom, J. J., Quist, L. *et al.* (2003). Association of mannose-binding lectin polymorphisms with sepsis and fatal outcome, in patients with systemic inflammatory response syndrome. *J Infect Dis*, **188**, 1394–403.

Geijtenbeek, T. B., Kwon, D. S., Torensma, R. *et al.* (2000*a*). DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell*, **100**, 587–97.

Geijtenbeek, T. B., Torensma, R., van Vliet, S. J. *et al.* (2000*b*). Identification of DC-SIGN, a novel dendritic cell-specific ICAM-3 receptor that supports primary immune responses. *Cell*, **100**, 575–85.

Goldfeld, A. E., Delgado, J. C., Thim, S. *et al.* (1998). Association of an *HLA-DQ* allele with clinical tuberculosis. *JAMA*, **279**, 226–8.

Goldgaber, D., Goldfarb, L. G., Brown, P. *et al.* (1989). Mutations in familial Creutzfeldt-Jakob disease and Gerstmann–Straussler–Scheinker's syndrome. *Exp Neurol*, **106**, 204–6.

Gomi, K., Kawasaki, K., Kawai, Y., Shiozaki, M. and Nishijima, M. (2002). Toll-like receptor 4-MD-2 complex mediates the signal transduction induced by flavolipin, an amino acid-containing lipid unique to *Flavobacterium meningosepticum*. *J Immunol*, **168**, 2939–43.

Goueffon, S. and du Saussay, C. (1969). Systematic survey on hemoglobin E and glucose-6-phosphate dehydrogenase in Cambodia (October 1965–June 1966). *Bull Soc Pathol Exot Filiales*, **62**, 1118–32.

Grange, J. M., Davies, P. D., Brown, R. C., Woodhead, J. S. and Kardjito, T. (1985). A study of vitamin D levels in Indonesian patients with untreated pulmonary tuberculosis. *Tubercle*, **66**, 187–91.

Grau, G. E., Taylor, T. E., Molyneux, M. E. *et al.* (1989). Tumor necrosis factor and disease severity in children with falciparum malaria. *N Engl J Med*, **320**, 1586–91.

Greenwood, C. M., Fujiwara, T. M., Boothroyd, L. J. *et al.* (2000). Linkage of tuberculosis to chromosome 2q35 loci, including NRAMP1, in a large aboriginal Canadian family. *Am J Hum Genet*, **67**, 405–16.

Greiner, J., Schleiermacher, E., Smith, T., Lenhard, V. and Vogel, F. (1978). The HLA system and leprosy in Thailand. *Hum Genet*, **42**, 201–13.

Gros, P., Skamene, E. and Forget, A. (1981). Genetic control of natural resistance to *Mycobacterium bovis* (BCG) in mice. *J Immunol*, **127**, 2417–21.

Haile, R. W., Iselius, L., Fine, P. E. and Morton, N. E. (1985). Segregation and linkage analyses of 72 leprosy pedigrees. *Hum Hered*, **35**, 43–52.

Hajjar, A.M., O'Mahony, D.S., Ozinsky, A. *et al.* (2001). Cutting edge: functional interactions between toll-like receptor (TLR) 2 and TLR1 or TLR6 in response to phenol-soluble modulin. *J Immunol*, **166**, 15–19.

Haldane, J.B.S. (1949). Disease and evolution. *Ric Sci*, Suppl A(**19**), 68–76.

Hamajima, N., Matsuo, K., Saito, T. *et al.* (2001*a*). Interleukin 1 polymorphisms, lifestyle factors, and *Helicobacter pylori* infection. *Jpn J Cancer Res*, **92**, 383–9.

Hamajima, N., Matsuo, K., Suzuki, T. *et al.* (2001*b*). Low expression myeloperoxidase genotype negatively associated with *Helicobacter pylori* infection. *Jpn J Cancer Res*, **92**, 488–93.

Hamblin, M.T. and Di Rienzo, A. (2000). Detection of the signature of natural selection in humans: evidence from the Duffy blood group locus. *Am J Hum Genet*, **66**, 1669–79.

Hashimoto, C., Hudson, K.L. and Anderson, K.V. (1988). The Toll gene of *Drosophila*, required for dorsal-ventral embryonic polarity, appears to encode a transmembrane protein. *Cell*, **52**, 269–79.

Hayashi, F., Smith, K.D., Ozinsky, A. *et al.* (2001). The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature*, **410**, 1099–103.

Heil, F., Ahmad-Nejad, P., Hemmi, H. *et al.* (2003). The Toll-like receptor 7 (TLR7)-specific stimulus loxoribine uncovers a strong relationship within the TLR7, 8 and 9 subfamily. *Eur J Immunol*, **33**, 2987–97.

Heinzmann, A., Ahlert, I., Kurz, T., Berner, R. and Deichmann, K.A. (2004). Association study suggests opposite effects of polymorphisms within IL8 on bronchial asthma and respiratory syncytial virus bronchiolitis. *J Allergy Clin Immunol*, **114**, 671–6.

Hemmi, H., Kaisho, T., Takeuchi, O. *et al.* (2002). Small anti-viral compounds activate immune cells via the TLR7 MyD88-dependent signaling pathway. *Nat Immunol*, **3**, 196–200.

Hemmi, H., Takeuchi, O., Kawai, T. *et al.* (2000). A Toll-like receptor recognizes bacterial DNA. *Nature*, **408**, 740–5.

Hendel, H., Caillat-Zucman, S., Lebuanec, H. *et al.* (1999). New class I and II HLA alleles strongly associated with opposite patterns of progression to AIDS. *J Immunol*, **162**, 6942–6.

Hendriks, J.C., Medley, G.F., van Griensven, G.J. *et al.* (1993). The treatment-free incubation period of AIDS in a cohort of homosexual men. *AIDS*, **7**, 231–9.

Henneke, P., Takeuchi, O., van Strijp, J.A. *et al.* (2001). Novel engagement of CD14 and multiple toll-like receptors by group B streptococci. *J Immunol*, **167**, 7069–76.

Hennig, B.J., Hellier, S., Frodsham, A.J. *et al.* (2002). Association of low-density lipoprotein receptor polymorphisms and outcome of hepatitis C infection. *Genes Immun*, **3**, 359–67.

Hill, A.V., Allsopp, C.E., Kwiatkowski, D. *et al.* (1991). Common west African HLA antigens are associated with protection from severe malaria. *Nature*, **352**, 595–600.

Hill, A.V., Whitehouse, D.B., Bowden, D.K. *et al.* (1987). Ahaptoglobinaemia in Melanesia: DNA and malarial antibody studies. *Trans R Soc Trop Med Hyg*, **81**, 573–7.

Hobbs, M.R., Udhayakumar, V., Levesque, M.C. *et al.* (2002). A new NOS2 promoter polymorphism associated with increased nitric oxide production and protection from severe malaria in Tanzanian and Kenyan children. *Lancet*, **360**, 1468–75.

Hoebee, B., Rietveld, E., Bont, L. *et al.* (2003). Association of severe respiratory syncytial virus bronchiolitis with interleukin-4 and interleukin-4 receptor alpha polymorphisms. *J Infect Dis*, **187**, 2–11.

Hoshino, K., Takeuchi, O., Kawai, T. *et al.* (1999). Cutting edge: Toll-like receptor 4 (TLR4)-deficient mice are hyporesponsive to lipopolysaccharide: evidence for TLR4 as the *Lps* gene product. *J Immunol*, **162**, 3749–52.

Huang, Y., Paxton, W.A., Wolinsky, S.M. *et al.* (1996). The role of a mutant CCR5 allele in HIV-1 transmission and disease progression. *Nat Med*, **2**, 1240–3.

Hull, J., Ackerman, H., Isles, K. *et al.* (2001). Unusual haplotypic structure of IL8, a susceptibility locus for a common respiratory virus. *Am J Hum Genet*, **69**, 413–19.

Hull, J., Rowlands, K., Lockhart, E. *et al.* (2003). Variants of the chemokine receptor CCR5 are associated with severe bronchiolitis caused by respiratory syncytial virus. *J Infect Dis*, **188**, 904–7.

Hull, J., Rowlands, K., Lockhart, E. *et al.* (2004). Haplotype mapping of the bronchiolitis susceptibility locus near IL8. *Hum Genet*, **114**, 272–9.

Hutagalung, R., Wilairatana, P., Looareesuwan, S. *et al.* (1999). Influence of hemoglobin E trait on the severity of Falciparum malaria. *J Infect Dis*, **179**, 283–6.

Jarolim, P., Palek, J., Amato, D. *et al.* (1991). Deletion in erythrocyte band 3 gene in malaria-resistant Southeast Asian ovalocytosis. *Proc Natl Acad Sci USA*, **88**, 11022–6.

Jeannet, M., Sztajzel, R., Carpentier, N., Hirschel, B. and Tiercy, J.M. (1989). HLA antigens are risk factors for

development of AIDS. *J Acquir Immune Defic Syndr*, **2**, 28–32.

Jeannin, P., Renno, T., Goetsch, L. *et al.* (2000). OmpA targets dendritic cells, induces their maturation and delivers antigen into the MHC class I presentation pathway. *Nat Immunol*, **1**, 502–9.

Jepson, A., Banya, W., Hassan-King, M. *et al.* (1994). Twin children in The Gambia: evidence for genetic regulation of physical characteristics in the presence of sub-optimal nutrition. *Ann Trop Paediatr*, **14**, 309–13.

Jepson, A., Fowler, A., Banya, W. *et al.* (2001). Genetic regulation of acquired immune responses to antigens of *mycobacterium tuberculosis*: a study of twins in West Africa. *Infect Immun*, **69**, 3989–94.

Jouanguy, E., Altare, F., Lamhamedi, S. *et al.* (1996). Interferon-gamma-receptor deficiency in an infant with fatal bacille Calmette-Guerin infection. *N Engl J Med*, **335**, 1956–61.

Kallmann, F. J. and Reisner, D. (1942). Twin studies on the significance of genetic factors in tuberculosis. *Am Rev Respir Dis*, **479**, 549–74.

Kang, T. J., Lee, S. B. and Chae, G. T. (2002). A polymorphism in the toll-like receptor 2 is associated with IL-12 production from monocyte in lepromatous leprosy. *Cytokine*, **20**, 56–62.

Karplus, T. M., Jeronimo, S. M., Chang, H. *et al.* (2002). Association between the tumor necrosis factor locus and the clinical outcome of *Leishmania chagasi* infection. *Infect Immun*, **70**, 6919–25.

Kaslow, R. A., Carrington, M., Apple, R. *et al.* (1996). Influence of combinations of human major histocompatibility complex genes on the course of HIV-1 infection. *Nat Med*, **2**, 405–11.

Kaur, G., Sachdeva, G., Bhutani, L. K. and Bamezai, R. (1997). Association of polymorphism at COL3A and CTLA4 loci on chromosome 2q31–33 with the clinical phenotype and in-vitro CMI status in healthy and leprosy subjects: a preliminary study. *Hum Genet*, **100**, 43–50.

Keet, I. P., Tang, J., Klein, M. R. *et al.* (1999). Consistent associations of HLA class I and II and transporter gene products with progression of human immunodeficiency virus type 1 infection in homosexual men. *J Infect Dis*, **180**, 299–309.

Kim, H. S., Park, M. H., Song, E. Y. *et al.* (2005). Association of HLA-DR and HLA-DQ genes with susceptibility to pulmonary tuberculosis in Koreans: preliminary evidence of associations with drug resistance, disease severity, and disease recurrence. *Hum Immunol*, **66**, 1074–81.

Knight, J. C., Udalova, I., Hill, A. V. *et al.* (1999). A polymorphism that affects OCT-1 binding to the TNF promoter region is associated with severe malaria. *Nat Genet*, **22**, 145–50.

Koch, O., Awomoyi, A., Usen, S. *et al.* (2002). IFNGR1 gene promoter polymorphisms and susceptibility to cerebral malaria. *J Infect Dis*, **185**, 1684–7.

Koppel, E. A., Saeland, E., de Cooker, D. J., van Kooyk, Y. and Geijtenbeek, T. B. (2005). DC-SIGN specifically recognizes Streptococcus pneumoniae serotypes 3 and 14. *Immunobiology*, **210**, 203–10.

Kostrikis, L. G., Huang, Y., Moore, J. P. *et al.* (1998). A chemokine receptor CCR2 allele delays HIV-1 disease progression and is associated with a CCR5 promoter mutation. *Nat Med*, **4**, 350–3.

Kouriba, B., Chevillard, C., Bream, J. H. *et al.* (2005). Analysis of the 5q31-q33 locus shows an association between IL13-1055C/T IL-13-591A/G polymorphisms and Schistosoma haematobium infections. *J Immunol*, **174**, 6274–81.

Krutzik, S. R., Ochoa, M. T., Sieling, P. A. *et al.* (2003). Activation and regulation of Toll-like receptors 2 and 1 in human leprosy. *Nat Med*, **9**, 525–32.

Kun, J. F., Mordmuller, B., Lell, B. *et al.* (1998). Polymorphism in promoter region of inducible nitric oxide synthase gene and protection against malaria. *Lancet*, **351**, 265–6.

Kun, J. F., Mordmuller, B., Perkins, D. J. *et al.* (2001). Nitric oxide synthase 2 (Lambarene) (G-954C), increased nitric oxide production, and protection against malaria. *J Infect Dis*, **184**, 330–6.

Kuroda, Y., Kaneoka, H., Shibasaki, H., Kume, S. and Yamaguchi, M. (1986). HLA study of Japanese patients with Creutzfeldt-Jakob disease: significant association with HLA-DQw3. *Ann Neurol*, **20**, 356–9.

Kurt-Jones, E. A., Popova, L., Kwinn, L. *et al.* (2000). Pattern recognition receptors TLR4 and CD14 mediate response to respiratory syncytial virus. *Nat Immunol*, **1**, 398–401.

Levin, M., Newport, M. J., D'Souza, S. *et al.* (1995). Familial disseminated atypical mycobacterial infection in childhood: a human mycobacterial susceptibility gene? *Lancet*, **345**, 79–83.

Li, C. M., Campbell, S. J., Kumararatne, D. S. *et al.* (2002). Association of a polymorphism in the *P2X7* gene with tuberculosis in a Gambian population. *J Infect Dis*, **186**, 1458–62.

Lin, T.M., Chen, C.J., Wu, M.M. *et al.* (1989). Hepatitis B virus markers in Chinese twins. *Anticancer Res*, **9**, 737–41.

Lindesmith, L., Moe, C., Marionneau, S. *et al.* (2003). Human susceptibility and resistance to Norwalk virus infection. *Nat Med*, **9**, 548–53.

Lio, D., Caruso, C., Di Stefano, R. *et al.* (2003). IL-10 and TNF-alpha polymorphisms and the recovery from HCV infection. *Hum Immunol*, **64**, 674–80.

Lio, D., Marino, V., Serauto, A. *et al.* (2002). Genotype frequencies of the +874T–>A single nucleotide polymorphism in the first intron of the interferon-gamma gene in a sample of Sicilian patients affected by tuberculosis. *Eur J Immunogenet*, **29**, 371–4.

Liu, R., Paxton, W.A., Choe, S. *et al.* (1996). Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell*, **86**, 367–77.

Lopez-Maderuelo, D., Arnalich, F., Serantes, R. *et al.* (2003). Interferon-gamma and interleukin-10 gene polymorphisms in pulmonary tuberculosis. *Am J Respir Crit Care Med*, **167**, 970–5.

Luoni, G., Verra, F., Arca, B. *et al.* (2001). Antimalarial antibody levels and IL-4 polymorphism in the Fulani of West Africa. *Genes Immun*, **2**, 411–14.

Maas, J., de Roda Husman, A.M., Brouwer, M. *et al.* (1998). Presence of the variant mannose-binding lectin alleles associated with slower progression to AIDS. Amsterdam Cohort Study. *Aids*, **12**, 2275–80.

Maier, A.G., Duraisingh, M.T., Reeder, J.C. *et al.* (2003). Plasmodium falciparum erythrocyte invasion through glycophorin C and selection for Gerbich negativity in human populations. *Nat Med*, **9**, 87–92.

Malik, S., Abel, L., Tooker, H. *et al.* (2005). Alleles of the *NRAMP1* gene are risk factors for pediatric tuberculosis disease. *Proc Natl Acad Sci USA*, **102**, 12183–8.

Malik, S., Arias, M., Di Flumeri, C., Garcia, L.F. and Schurr, E. (2003). Absence of association between mannose-binding lectin gene polymorphisms and HIV-1 infection in a Colombian population. *Immunogenetics*, **55**, 49–52.

Malik, S., Greenwood, C.M., Eguale, T. *et al.* (2006). Variants of the *SFTPA1* and *SFTPA2* genes and susceptibility to tuberculosis in Ethiopia. *Hum Genet*, **118**, 752–9.

Marquet, S., Abel, L., Hillaire, D. *et al.* (1996). Genetic localization of a locus controlling the intensity of infection by *Schistosoma mansoni* on chromosome 5q31-q33. *Nat Genet*, **14**, 181–4.

Martin, M.P., Dean, M., Smith, M.W. *et al.* (1998). Genetic acceleration of AIDS progression by a promoter variant of CCR5. *Science*, **282**, 1907–11.

Massari, P., Henneke, P., Ho, Y. *et al.* (2002). Cutting edge: immune stimulation by neisserial porins is Toll-like receptor 2 and MyD88 dependent. *J Immunol*, **168**, 1533–7.

McGuire, W., Hill, A.V., Allsopp, C.E., Greenwood, B.M. and Kwiatkowski, D. (1994). Variation in the TNF-alpha promoter region associated with susceptibility to cerebral malaria. *Nature*, **371**, 508–10.

McGuire, W., Knight, J.C., Hill, A.V. *et al.* (1999). Severe malarial anemia and cerebral malaria are associated with different tumor necrosis factor promoter alleles. *J Infect Dis*, **179**, 287–90.

Mead, S., Mahal, S.P., Beck, J. *et al.* (2001). Sporadic – but not variant – Creutzfeldt-Jakob disease is associated with polymorphisms upstream of PRNP exon 1. *Am J Hum Genet*, **69**, 1225–35.

Mead, S., Stumpf, M.P., Whitfield, J. *et al.* (2003). Balancing selection at the prion protein gene consistent with prehistoric kurulike epidemics. *Science*, **300**, 640–3.

Means, T.K., Wang, S., Lien, E. *et al.* (1999). Human Toll-like receptors mediate cellular activation by *Mycobacterium tuberculosis*. *J Immunol*, **163**, 3920–7.

Mehra, N.K., Rajalingam, R., Mitra, D.K., Taneja, V. and Giphart, M.J. (1995). Variants of HLA-DR2/DR51 group haplotypes and susceptibility to tuberculoid leprosy and pulmonary tuberculosis in Asian Indians. *Int J Lepr Other Mycobact Dis*, **63**, 241–8.

Meisner, S.J., Mucklow, S., Warner, G., Sow, S.O., Lienhardt, C. and Hill, A.V. (2001). Association of NRAMP1 polymorphism with leprosy type but not susceptibility to leprosy per se in west Africans. *Am J Trop Med Hyg*, **65**, 733–5.

Mgone, C.S., Koki, G., Paniu, M.M. *et al.* (1996). Occurrence of the erythrocyte band 3 (AE1) gene deletion in relation to malaria endemicity in Papua New Guinea. *Trans R Soc Trop Med Hyg*, **90**, 228–31.

Michaud, C.M., Murray, C.J. and Bloom, B.R. (2001). Burden of disease – implications for future research. *JAMA*, **285**, 535–9.

Migueles, S.A., Sabbaghian, M.S., Shupert, W.L. *et al.* (2000). HLA B*5701 is highly associated with restriction of virus replication in a subgroup of HIV-infected long term nonprogressors. *Proc Natl Acad Sci USA*, **97**, 2709–14.

Miller, L.H., Mason, S.J., Dvorak, J.A., McGinniss, M.H. and Rothman, I.K. (1975). Erythrocyte receptors for

(*Plasmodium knowlesi*) malaria: Duffy blood group determinants. *Science*, **189**, 561–3.

Mira, J. P., Cariou, A., Grall, F. *et al.* (1999). Association of TNF2, a TNF-alpha promoter polymorphism, with septic shock susceptibility and mortality: a multicenter study. *JAMA*, **282**, 561–8.

Mira, M. T., Alcais, A., Nguyen, V. T. *et al.* (2004). Susceptibility to leprosy is associated with PARK2 and PACRG. *Nature*, **427**, 636–40.

Mira, M. T., Alcais, A., Van Thuc, N. *et al.* (2003). Chromosome 6q25 is linked to susceptibility to leprosy in a Vietnamese population. *Nat Genet*, **33**, 412–15.

Miyanaga, K., Juji, T., Maeda, H., Nakajima, S. and Kobayashi, S. (1981). Tuberculoid leprosy and HLA in Japanese. *Tissue Antigens*, **18**, 331–4.

Modiano, D., Luoni, G., Sirima, B. S. *et al.* (2001). Haemoglobin C protects against clinical Plasmodium falciparum malaria. *Nature*, **414**, 305–8.

Modiano, D., Petrarca, V., Sirima, B. S. *et al.* (1996). Different response to Plasmodium falciparum malaria in west African sympatric ethnic groups. *Proc Natl Acad Sci USA*, **93**, 13206–11.

Mohamed, H. S., Ibrahim, M. E., Miller, E. N. *et al.* (2003). Genetic susceptibility to visceral leishmaniasis in the Sudan: linkage and association with IL4 and IFNGR1. *Genes Immun*, **4**, 351–5.

Mohamed, H. S., Ibrahim, M. E., Miller, E. N. *et al.* (2004). *SLC11A1* (formerly *NRAMP1*) and susceptibility to visceral leishmaniasis in the Sudan. *Eur J Hum Genet*, **12**, 66–74.

Mummidi, S., Ahuja, S. S., Gonzalez, E. *et al.* (1998). Genealogy of the CCR5 locus and chemokine system gene variants associated with altered rates of HIV-1 disease progression. *Nat Med*, **4**, 786–93.

Nagase, H., Okugawa, S., Ota, Y. *et al.* (2003). Expression and function of Toll-like receptors in eosinophils: activation by Toll-like receptor 7 ligand. *J Immunol*, **171**, 3977–82.

Nakayama, E. E., Hoshino, Y., Xin, X. *et al.* (2000). Polymorphism in the interleukin-4 promoter affects acquisition of human immunodeficiency virus type 1 syncytium-inducing phenotype. *J Virol*, **74**, 5452–9.

Newport, M. J., Huxley, C. M., Huston, S. *et al.* (1996). A mutation in the interferon-gamma-receptor gene and susceptibility to mycobacterial infection. *N Engl J Med*, **335**, 1941–9.

Ohashi, K., Burkart, V., Flohe, S. and Kolb, H. (2000). Cutting edge: heat shock protein 60 is a putative endogenous ligand of the Toll-like receptor-4 complex. *J Immunol*, **164**, 558–61.

Omi, K., Ohashi, J., Patarapotikul, J. *et al.* (2002). Fcgamma receptor IIA and IIIB polymorphisms are associated with susceptibility to cerebral malaria. *Parasitol Int*, **51**, 361–6.

Omi, K., Ohashi, J., Patarapotikul, J. *et al.* (2003). CD36 polymorphism is associated with protection from cerebral malaria. *Am J Hum Genet*, **72**, 364–74.

Opitz, B., Schroder, N. W., Spreitzer, I. *et al.* (2001). Toll-like receptor-2 mediates Treponema glycolipid and lipoteichoic acid-induced NF-kappaB translocation. *J Biol Chem*, **276**, 22041–7.

Owen, F., Poulter, M., Lofthouse, R. *et al.* (1989). Insertion in prion protein gene in familial Creutzfeldt-Jakob disease. *Lancet*, **1**, 51–2.

Ozinsky, A., Underhill, D. M., Fontenot, J. D. *et al.* (2000). The repertoire for pattern recognition of pathogens by the innate immune system is defined by cooperation between Toll-like receptors. *Proc Natl Acad Sci USA*, **97**, 13766–71.

Patel, S. S., Mehlotra, R. K., Kastens, W. *et al.* (2001). The association of the glycophorin C exon 3 deletion with ovalocytosis and malaria susceptibility in the Wosera, Papua New Guinea. *Blood*, **98**, 3489–91.

Petersen, D. C., Glashoff, R. H., Shrestha, S. *et al.* (2005). Risk for HIV-1 infection associated with a common CXCL12 (SDF1) polymorphism and CXCR4 variation in an African population. *J Acquir Immune Defic Syndr*, **40**, 521–6.

Picard, C., Fieschi, C., Altare, F. *et al.* (2002). Inherited interleukin-12 deficiency: IL12B genotype and clinical phenotype of 13 patients from six kindreds. *Am J Hum Genet*, **70**, 336–48.

Poltorak, A., He, X., Smirnova, I. *et al.* (1998*a*). Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in *Tlr4* gene. *Science*, **282**, 2085–8.

Poltorak, A., Smirnova, I., He, X. *et al.* (1998*b*). Genetic and physical mapping of the Lps locus: identification of the toll-4 receptor as a candidate gene in the critical region. *Blood Cells Mol Dis*, **24**, 340–55.

Pravica, V., Perrey, C., Stevens, A., Lee, J. H. and Hutchinson, I. V. (2000). A single nucleotide polymorphism in the first intron of the human IFN-gamma gene: absolute correlation with a polymorphic CA microsatellite marker of high IFN-gamma production. *Hum Immunol*, **61**, 863–6.

Puthothu, B., Krueger, M., Heinze, J., Forster, J. and Heinzmann, A. (2006). Impact of IL8 and

IL8-Receptor alpha polymorphisms on the genetics of bronchial asthma and severe RSV infections. *Clin Mol Allergy*, **4**, 2.

Quaye, I.K., Ekuban, F.A., Goka, B.Q. *et al.* (2000). Haptoglobin 1–1 is associated with susceptibility to severe *Plasmodium falciparum* malaria. *Trans R Soc Trop Med Hyg*, **94**, 216–19.

Rad, R., Prinz, C., Neu, B. *et al.* (2003). Synergistic effect of *Helicobacter pylori* virulence factors and interleukin-1 polymorphisms for the development of severe histological changes in the gastric mucosa. *J Infect Dis*, **188**, 272–81.

Rajalingam, R., Mehra, N.K. and Singal, D.P. (2000). Polymorphism in heat-shock protein 70-1 (HSP70-1) gene promoter region and susceptibility to tuberculoid leprosy and pulmonary tuberculosis in Asian Indians. *Indian J Exp Biol*, **38**, 658–62.

Rajalingam, R., Singal, D.P. and Mehra, N.K. (1997). Transporter associated with antigen-processing (*TAP*) genes and susceptibility to tuberculoid leprosy and pulmonary tuberculosis. *Tissue Antigens*, **49**, 168–72.

Ramaley, P.A., French, N., Kaleebu, P. *et al.* (2002). HIV in Africa (Communication arising): chemokine-receptor genes and AIDS risk. *Nature*, **417**, 140.

Rani, R., Zaheer, S.A. and Mukherjee, R. (1992). Do human leukocyte antigens have a role to play in differential manifestation of multibacillary leprosy: a study on multibacillary leprosy patients from north India. *Tissue Antigens*, **40**, 124–7.

Rea, T.H., Levan, N.E. and Terasaki, P.I. (1976). Histocompatibility antigens in patients with leprosy. *J Infect Dis*, **134**, 615–18.

Reiling, N., Holscher, C., Fehrenbach, A. *et al.* (2002). Cutting edge: Toll-like receptor (TLR)2- and TLR4-mediated pathogen recognition in resistance to airborne infection with *Mycobacterium tuberculosis*. *J Immunol*, **169**, 3480–4.

Rihet, P., Abel, L., Traore, Y. *et al.* (1998*a*). Human malaria: segregation analysis of blood infection levels in a suburban area and a rural area in Burkina Faso. *Genet Epidemiol*, **15**, 435–50.

Rihet, P., Traore, Y., Abel, L. *et al.* (1998*b*). Malaria in humans: *Plasmodium falciparum* blood infection levels are linked to chromosome 5q31-q33. *Am J Hum Genet*, **63**, 498–505.

Rock, F.L., Hardiman, G., Timans, J.C., Kastelein, R.A. and Bazan, J.F. (1998). A family of human receptors structurally related to *Drosophila Toll*. *Proc Natl Acad Sci USA*, **95**, 588–93.

Rougemont, A., Quilici, M., Delmont, J. and Ardissone, J.P. (1980). Is the HpO phenomenon in tropical populations really genetic? *Hum Hered*, **30**, 201–3.

Rossouw, M., Nel, H.J., Cooke, G.S., van Helden, P.D. and Hoal, E.G. (2003). Association between tuberculosis and a polymorphic NFkappaB binding site in the interferon gamma gene. *Lancet*, **361**, 1871–2.

Roy, S., Frodsham, A., Saha, B. *et al.* (1999). Association of vitamin D receptor genotype with leprosy type. *J Infect Dis*, **179**, 187–91.

Roy, S., McGuire, W., Mascie-Taylor, C.G. *et al.* (1997). Tumor necrosis factor promoter polymorphism and susceptibility to lepromatous leprosy. *J Infect Dis*, **176**, 530–2.

Ruwende, C. and Hill, A. (1998). Glucose-6-phosphate dehydrogenase deficiency and malaria. *J Mol Med*, **76**, 581–8.

Ruwende, C., Khoo, S.C., Snow, R.W. *et al.* (1995). Natural selection of hemi- and heterozygotes for G6PD deficiency in Africa by resistance to severe malaria. *Nature*, **376**, 246–9.

Sabeti, P.C., Reich, D.E., Higgins, J.M. *et al.* (2002). Detecting recent positive selection in the human genome from haplotype structure. *Nature*, **419**, 832–7.

Saiki, R.K., Scharf, S., Faloona, F. *et al.* (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, **230**, 1350–4.

Sakai, T., Matsuoka, M., Aoki, M., Nosaka, K. and Mitsuya, H. (2001). Missense mutation of the interleukin-12 receptor beta1 chain-encoding gene is associated with impaired immunity against *Mycobacterium avium* complex infection. *Blood*, **97**, 2688–94.

Sakuntabhai, A., Turbpaiboon, C., Casademont, I. *et al.* (2005). A variant in the CD209 promoter is associated with severity of dengue disease. *Nat Genet*, **37**, 507–13.

Salkowitz, J.R., Bruse, S.E., Meyerson, H. *et al.* (2003). CCR5 promoter polymorphism determines macrophage CCR5 density and magnitude of HIV-1 propagation in vitro. *Clin Immunol*, **108**, 234–40.

Samson, M., Libert, F., Doranz, B.J. *et al.* (1996). Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature*, **382**, 722–5.

Sanguansermsri, T., Flatz, S.D. and Flatz, G. (1987). The hemoglobin E belt at the Thailand–Kampuchea border: ethnic and environmental determinants of

hemoglobin E and beta-thalassemia gene frequencies. *Gene Geogr*, **1**, 155–61.

Saunders, M. A., Hammer, M. F. and Nachman, M. W. (2002). Nucleotide variability at G6pd and the signature of malarial selection in humans. *Genetics*, **162**, 1849–61.

Schauf, V., Ryan, S., Scollard, D. *et al.* (1985). Leprosy associated with HLA-DR2 and DQw1 in the population of northern Thailand. *Tissue Antigens*, **26**, 243–7.

Schroder, N. W., Morath, S., Alexander, C. *et al.* (2003). Lipoteichoic acid (LTA) of *Streptococcus pneumoniae* and *Staphylococcus aureus* activates immune cells via *Toll*-like receptor (TLR)-2, lipopolysaccharide-binding protein (LBP), and CD14, whereas TLR-4 and MD-2 are not involved. *J Biol Chem*, **278**, 15587–94.

Serjeantson, S. W. (1983). HLA and susceptibility to leprosy. *Immunol Rev*, **70**, 89–112.

Serrano-Gomez, D., Leal, J. A. and Corbi, A. L. (2005). DC-SIGN mediates the binding of *Aspergillus fumigatus* and keratinophylic fungi by human dendritic cells. *Immunobiology*, **210**, 175–83.

Shaffer, N., Grau, G. E., Hedberg, K. *et al.* (1991). Tumor necrosis factor and severe malaria. *J Infect Dis*, **163**, 96–101.

Shaw, M. A., Donaldson, I. J., Collins, A. *et al.* (2001). Association and linkage of leprosy phenotypes with HLA class II and tumour necrosis factor genes. *Genes Immun*, **2**, 196–204.

Shi, Y. P., Nahlen, B. L., Kariuki, S. *et al.* (2001). Fcgamma receptor IIa (CD32) polymorphism is associated with protection of infants against high-density *Plasmodium falciparum* infection. VII. Asembo Bay Cohort Project. *J Infect Dis*, **184**, 107–11.

Shin, H. D., Winkler, C., Stephens, J. C. *et al.* (2000). Genetic restriction of HIV-1 pathogenesis to AIDS by promoter alleles of IL10. *Proc Natl Acad Sci USA*, **97**, 14467–72.

Siddiqui, M. R., Meisner, S., Tosh, K. *et al.* (2001). A major susceptibility locus for leprosy in India maps to chromosome 10p13. *Nat Genet*, **27**, 439–41.

Singh, S. P., Mehra, N. K., Dingley, H. B., Pande, J. N. and Vaidya, M. C. (1983). Human leukocyte antigen (HLA)-linked control of susceptibility to pulmonary tuberculosis and association with HLA-DR types. *J Infect Dis*, **148**, 676–81.

Sjoberg, K., Lepers, J. P., Raharimalala, L. *et al.* (1992). Genetic regulation of human anti-malarial antibodies in twins. *Proc Natl Acad Sci USA*, **89**, 2101–4.

Smith, M. W., Dean, M., Carrington, M. *et al.* (1997). Contrasting genetic influence of CCR2 and CCR5 variants on HIV-1 infection and disease progression. Hemophilia Growth and Development Study (HGDS), Multicenter AIDS Cohort Study (MACS), Multicenter Hemophilia Cohort Study (MHCS), San Francisco City Cohort (SFCC), ALIVE Study. *Science*, **277**, 959–65.

Smolnikova, M. V. and Konenkov, V. I. (2002). Association of IL2, TNFA, IL4 and IL10 promoter gene polymorphisms with the rate of progression of the HIV infection. *Russ J Immunol*, **7**, 349–56.

Sorensen, T. I., Nielsen, G. G., Andersen, P. K. and Teasdale, T. W. (1988). Genetic and environmental influences on premature death in adult adoptees. *N Engl J Med*, **318**, 727–32.

Takeuchi, O., Kawai, T., Muhlradt, P. F. *et al.* (2001). Discrimination of bacterial lipoproteins by *Toll*-like receptor 6. *Int Immunol*, **13**, 933–40.

Tal, G., Mandelberg, A., Dalal, I. *et al.* (2004). Association between common *Toll*-like receptor 4 mutations and severe respiratory syncytial virus disease. *J Infect Dis*, **189**, 2057–63.

Tang, G. J., Huang, S. L., Yien, H. W. *et al.* (2000). Tumor necrosis factor gene polymorphism and septic shock in surgical infection. *Crit Care Med*, **28**, 2733–6.

Tang, J., Costello, C., Keet, I. P. *et al.* (1999). HLA class I homozygosity accelerates disease progression in human immunodeficiency virus type 1 infection. *AIDS Res Hum Retroviruses*, **15**, 317–24.

Thursz, M. R., Kwiatkowski, D., Allsopp, C. E. *et al.* (1995). Association between an MHC class II allele and clearance of hepatitis B virus in The Gambia. *N Engl J Med*, **332**, 1065–9.

Tomiyama, H., Miwa, K., Shiga, H. *et al.* (1997). Evidence of presentation of multiple HIV-1 cytotoxic T lymphocyte epitopes by HLA-B*3501 molecules that are associated with the accelerated progression of AIDS. *J Immunol*, **158**, 5026–34.

Tosh, K., Meisner, S., Siddiqui, M. R. *et al.* (2002). A region of chromosome 20 is linked to leprosy susceptibility in a South Indian population. *J Infect Dis*, **186**, 1190–3.

Tournamille, C., Colin, Y., Cartron, J. P. and Le Van Kim, C. (1995). Disruption of a GATA motif in the Duffy gene promoter abolishes erythroid gene expression in Duffy-negative individuals. *Nat Genet*, **10**, 224–8.

Trembath, R. C., Clough, R. L., Rosbotham, J. L. *et al.* (1997). Identification of a major susceptibility locus on chromosome 6p and evidence for further disease loci

revealed by a two stage genome-wide search in psoriasis. *Hum Mol Genet*, **6**, 813–20.

Tso, H. W., Ip, W. K., Chong, W. P. *et al.* (2005). Association of interferon gamma and interleukin 10 genes with tuberculosis in Hong Kong Chinese. *Genes Immun*, **6**, 358–63.

Underhill, D. M., Ozinsky, A., Hajjar, A. M. *et al.* (1999a). The Toll-like receptor 2 is recruited to macrophage phagosomes and discriminates between pathogens. *Nature*, **401**, 811–15.

Underhill, D. M., Ozinsky, A., Smith, K. D. and Aderem, A. (1999*b*). Toll-like receptor-2 mediates mycobacteria-induced proinflammatory signaling in macrophages. *Proc Natl Acad Sci USA*, **96**, 14459–63.

Van Everbroeck, B., Croes, E. A., Pals, P. *et al.* (2001). Influence of the prion protein and the apolipoprotein E genotype on the Creutzfeldt-Jakob Disease phenotype. *Neurosci Lett*, **313**, 69–72.

Verhagen, C. E., de Boer, T., Smits, H. H. *et al.* (2000). Residual type 1 immunity in patients genetically deficient for interleukin 12 receptor beta1 (IL-12Rbeta1): evidence for an IL-12Rbeta1-independent pathway of IL-12 responsiveness in human T cells. *J Exp Med*, **192**, 517–28.

Vidal, S. M., Malo, D., Vogan, K., Skamene, E. and Gros, P. (1993). Natural resistance to infection with intracellular parasites: isolation of a candidate for Bcg. *Cell*, **73**, 469–85.

Wagener, D. K., Schauf, V., Nelson, K. E. *et al.* (1988). Segregation analysis of leprosy in families of northern Thailand. *Genet Epidemiol*, **5**, 95–105.

Wallace, C., Clayton, D. and Fine, P. (2003). Estimating the relative recurrence risk ratio for leprosy in Karonga District, Malawi. *Lepr Rev*, **74**, 133–40.

Wang, L. M., Kimura, A., Satoh, M. and Mineshita, S. (1999). HLA linked with leprosy in southern China: HLA-linked resistance alleles to leprosy. *Int J Lepr Other Mycobact Dis*, **67**, 403–8.

Wibawa, T., Soebono, H. and Matsuo, M. (2002). Association of a missense mutation of the laminin alpha2 gene with tuberculoid type of leprosy in Indonesian patients. *Trop Med Int Health*, **7**, 631–6.

Wilkinson, R. J., Llewelyn, M., Toossi, Z. *et al.* (2000). Influence of vitamin D deficiency and vitamin D receptor polymorphisms on tuberculosis among Gujarati Asians in west London: a case-control study. (See comments.) Lancet, **355**, 618–21.

Williams, T. N., Wambua, S., Uyoga, S. *et al.* (2005). Both heterozygous and homozygous {alpha}+ thalassemia protect against severe and fatal *Plasmodium falciparum* malaria on the coast of Kenya. *Blood*, **106**, 368–71.

Wilson, J., Rowlands, K., Rockett, K. *et al.* (2005). Genetic variation at the IL10 gene locus is associated with severity of respiratory syncytial virus bronchiolitis. *J Infect Dis*, **191**, 1705–9.

Winkler, C., Modi, W., Smith, M. W. *et al.* (1998). Genetic restriction of AIDS pathogenesis by an SDF-1 chemokine gene variant. ALIVE Study, Hemophilia Growth and Development Study (HGDS), Multicenter AIDS Cohort Study (MACS), Multicenter Hemophilia Cohort Study (MHCS), San Francisco City Cohort (SFCC). *Science*, **279**, 389–93.

Wyllie, D. H., Kiss-Toth, E., Visintin, A. *et al.* (2000). Evidence for an accessory protein function for *Toll*-like receptor 1 in anti-bacterial responses. *J Immunol*, **165**, 7125–32.

Yoon, S. K., Han, J. Y., Pyo, C. W. *et al.* (2005). Association between human leukocytes antigen alleles and chronic hepatitis C virus infection in the Korean population. *Liver Int*, **25**, 1122–7.

Yoshimura, A., Lien, E., Ingalls, R. R. *et al.* (1999). Cutting edge: recognition of Gram-positive bacterial cell wall components by the innate immune system occurs via *Toll*-like receptor 2. *J Immunol*, **163**, 1–5.

Yoshizawa, K., Ota, M., Saito, S. *et al.* (2003). Long-term follow-up of hepatitis C virus infection: HLA class II loci influences the natural history of the disease. *Tissue Antigens*, **61**, 159–65.

20

Inflammatory bowel diseases

Jean-Pierre Hugot

Introduction

Inflammatory bowel diseases (IBD) consist of two major disorders: Crohn's disease (CD, OMIM 266600) and ulcerative colitis (UC, OMIM 191390). They are both characterized by a chronic or relapsing inflammation of the digestive tract (for review see Shanahan, 2002; Podolsky, 2002). In UC, the inflammation is limited to the colon with continuous mucosal inflammation already affecting the rectum. On the other hand, CD may affect all the digestive tract from the mouth to the anus with discontinuous lesions. The inflammation is often transmural with potential complications including fistulas, abscesses and strictures. At late stages, granulomas with giant and epithelioid cells are encountered in biopsies or specimens in about half of CD cases.

UC and CD are usually diagnosed in patients presenting with isolated or associated symptoms such as: diarrhea, rectal bleeding, abdominal pain, inflammatory syndromes and malabsorption. Both disorders can be complicated by under-nutrition (and failure to grow in children), osteopenia, extra-intestinal inflammation and cancer. IBD treatment is often complex and requires a combination of anti-inflammatory drugs including 5-aminosalicylates and steroids, immunosuppressant agents and biological therapies. Surgery is often mandatory and iatrogenic complications are frequent.

IBD are lifelong disorders occurring in the young adult with a peak of incidence in the third decade (for review see Mayberry and Rhodes, 1984). CD is more frequent in females (M/F sex ratio = 0.8) while UC is more frequent in males (M/F sex ratio: 1.2). In the 1990s the mean annual incidence rate of UC ($10.4/10^5$ (CI: 7.6–13.1)) was higher than for CD ($5.6/10^5$ (CI: 2.8–8.3)) in Europe (Shivananda *et al.*, 1996). IBD prevalence is less precisely known but it can be roughly estimated that 2/1000 inhabitants have either CD or UC in occidental countries (Mayberry and Rhodes, 1984). Finally, the mortality directly attributable to IBD is low.

IBD are not evenly distributed around the world (for review see Mayberry and Rhodes, 1984). They are uncommon in developing areas. IBD are certainly under-diagnosed in many of these countries explaining in part the low incidence. Ethnic differences have also been suggested and early studies suggested a higher risk of IBD in White people. However, epidemiological data from recently emigrated people (Probert *et al.*, 1992) or recent analyses on different ethnic groups living in Western countries (Kugathasan *et al.*, 2003) have not supported a major role of ethnicity in IBD predisposition. On the contrary, recent reports of CD outbreaks in North Africa, South Asia or South America suggest that IBD incidence increases in parallel with the level of economic development.

Earlier reports from Western countries also suggest that IBD are related to the Western lifestyle. An increase in CD incidence was reported in nearly all of the West European and North American countries during the second part of the twentieth century (Brahme *et al.*, 1975; Evans and Acheson, 1965; Fellows *et al.*, 1990; Hellers, 1979; Kyle, 1971;

Table 20.1. Environmental factors investigated in IBD. Currently supported specific risk factors are in bold

Diet	Hygiene and comfort	Infectious agents	Others
Refined sugars	Hot tap water	*Mycobacterium paratuberculosis*	**Cigarette smoking**
Fast food and cola drinks	Perinatal infections	Measles	**Appendectomy**
Microparticles	Infections in childhood	Measles vaccination	Life events
Chewing-gum	Toothpaste	*E. Coli*	Oral contraceptives
Margarine	Breastfeeding	*Yersinia* species	
Fruit and vegetables	Adenoidectomy	*Listeria* monocytogenes	
Dairy products	Socioeconomic status	Antibiotics	
Baker's yeast	Soft toys		
Cornflakes	Refrigeration		
Coffee			
Curry			

Loftus *et al.*, 1998; Rose *et al.*, 1988). UC incidence also increased in some countries but to a lesser extent. The growth of CD seems to have occurred first in the United States, then in Northern Europe and later in Southern Europe. These differences may have in part contributed to the classic North–South gradient of incidence observed for CD (Shivananda *et al.*, 1996).

Even if some methodological biases can not definitively be ruled out, this secular trend in incidence strongly supports the effect of environmental factors. In the context of complex genetic traits such as IBD, it means that progressive exposure to the putative risk factors affected more and more genetically at-risk people. This point of view is consistent with the progressive lifestyle standardization in developed countries during the second half of the twentieth century. Furthermore, the present stabilization of the disease incidence in North America and in some European countries suggests that exposure to this risk factor is today widespread in Western countries and possibly that the majority of genetically at-risk people are now affected by the disease.

Environmental factors

If the long-term trend of CD demonstrates a role for the environment, epidemiological studies have failed to identify the major causative factor(s). Many environmental factors have been proposed as causative agents for IBD (Table 20.1). From these, diet, domestic hygiene and micro-organisms were the most investigated areas (Andersson *et al.*, 2001; Bull *et al.*, 2003; Cosnes *et al.*, 1999; Darfeuille-Michaud *et al.*, 1998; Ekbom *et al.*, 1990; Ekbom *et al.*, 1996; Gent *et al.*, 1994; Gilat *et al.*, 1987; Godet *et al.*, 1995; Greenstein, 2003; Kallinowski *et al.*, 1998; Lamps *et al.*, 2003; Persson *et al.*, 1987; Sanderson *et al.*, 1992; Thompson *et al.*, 1995). In fact the only risk factors that can be currently supported are appendectomy, which has a protective role in UC, and tobacco. Smoking protects against UC (OR = 0.41 (0.34–0.48) and improves the disease course (Calkins, 1989). On the other hand, smoking increases the risk of CD (OR = 2.0 (1.65–2.47)) and the severity of the disease, as shown by an increased use of immunosuppressant agents and surgical intervention in smokers (Calkins, 1989; Cosnes *et al.*, 1996; Cottone *et al.*, 1994; Lindberg *et al.*, 1992; Odes *et al.*, 2001; Sutherland *et al.*, 1990). Smoking cessation improves the disease course and it is thus a major goal of CD treatment (Cosnes *et al.*, 2001). The reason why tobacco influences IBD course is not clearly understood. Nicotine, carbon monoxide and other cigarette-derived substances regulate various functions related to the intestines, including mucus synthesis, cytokine production,

macrophage bactericidal properties, intestinal permeability to macromolecules, contractile activity and microcirculation. Finally, cigarette smoking may modulate the IBD phenotype, increasing the risk of CD in smokers and the risk of UC in non-smokers (Bridger *et al.*, 2002).

The resemblance between CD and infectious disorders such as intestinal tuberculosis or yersiniosis suggested that CD may be an infectious condition. The effect of bacterial colonization on the digestive tract in genetically engineered animal models and the clinical effect of some antibiotics and intestinal flux manipulation in the human disease reinforced the idea that IBD may be related to certain bacterial species. However, the main argument in favour of infectious agents in CD came from genetic studies showing an association between CD and *CARD15/NOD2* (see below).

Should IBD be an infectious condition, it is not highly contagious (according to the popular definition of the word) as demonstrated by the limited number of affected spouses (Laharie *et al.*, 2001) or medical care workers. As a result, IBD cannot be regarded as a classic infectious disease but it can be seen as a bacteria-related disorder. From this point of view, CD should be the result of an excessive inflammatory response of the host to an infectious agent present in the digestive tract. This over-reaction would be limited to genetically at-risk people, while the same infectious agent is likely to have only a modest pathogenic effect in non-genetically predisposed people. According to this hypothesis, CD can be related to common bacteria widely disseminated in the Western world.

The gut flora as a whole is usually proposed to be an etiological factor for IBD and these diseases are considered to be an abnormal response to the normal digestive flora. This opinion is mainly supported by the fact that *CARD15/NOD2* potentially recognizes every type of bacterium. However, several points do not fit this hypothesis well. Firstly, the gut flora hypothesis does not explain the long-term trend in incidence of the disease since this factor is unlikely to have changed in genetically at risk individuals. If the gut flora as a whole is a risk factor, a constant incidence of CD should be observed over time. It can be postulated that external environmental factors (for examples dietary change or antibiotic use) have altered the gut flora composition during the twentieth century. However, in that case, the entire gut flora can no longer be considered as a risk factor and the question turns toward more specific bacteria. Secondly, CD lesions may occur in some portions of the digestive tract which are nearly sterile (such as the stomach). Thirdly, the intracytoplasmic location of *CARD15* suggests that the infectious agents involved in CD are able to interact within the cytoplasm of the cell. Such bacteria can be intracytoplasmic or have a secretory apparatus. In both cases, they exhibit some invasive properties which are not observed for the usual saprophytes present in the gut.

As an alternative to the gut flora hypothesis, several specific viruses and bacteria have been studied, including species of *Mycobacterium, Listeria, Yersinia, Escherichia coli,* and Measles virus (Andersson *et al.*, 2001; Bull *et al.*, 2003; Darfeuille-Michaud *et al.*, 1998; Ekbom *et al.*, 1996; Greenstein, 2003; Kallinowski *et al.*, 1998; Lamps *et al.*, 2003; Sanderson *et al.*, 1992; Thompson *et al.*, 1995). However, none of them can be pinpointed as causal agents at the present time. But if an excessive inflammatory response to an infectious agent with a limited pathogenic effect is involved, it would be very difficult to identify it by conventional methods (as was the case for Whipple's disease). Few studies using modern methods for bacterial identification (such as Polymerase chain reaction (PCR) analyses of IBD biopsies) are currently available for IBD. Among the specific bacteria tested, culture isolation and PCR analyses were able to detect *Mycobacterium, Yersinia, Pseudomonas fluorescens* and *Escherichia coli* species in CD lesions (Bull *et al.*, 2003; Darfeuille-Michaud *et al.*, 1998; Kallinowski *et al.*, 1998; Lamps *et al.*, 2003). These specific bacteria thus appear to be the most relevant candidates today.

Genetic factors

The genetic component of IBD was initially suggested by early reports of familial aggregation of these diseases. In 1934, Crohn and colleagues reported the first familial aggregation of CD. In the 1950s and 1960s many other groups confirmed that IBD are more familial than expected by chance. On average, 6–8% of UC patients and 8–10% of CD patients have at least one affected relative (for review see Russell and Satsangi, 2004). Interestingly, values as high as 20% or more have been reported in children, suggesting that either pediatric CD is more genetic or that familial environmental risk factors are important. The non-random distribution of the affected siblings within sibships with multiple cases strongly argues for the second hypothesis of an environmental factor shared by nuclear family members (Hugot *et al.*, 2003*b*).

Pedigrees of IBD patients do not exhibit a Mendelian model of inheritance. The lifetime risk of developing IBD for first-degree relatives has been estimated from 1% to 10% with a lower risk for UC relatives (Russell and Satsangi, 2004). This large range of values is mainly explained by methodological differences between studies. Siblings of CD patients have the highest risk with an estimated sib risk ratio (λs) value from 25 to 42 (Bayless *et al.*, 1996; Monsen *et al.*, 1987; Peeters *et al.*, 1996; Roth *et al.*, 1989). The risk seems to decrease rapidly for more distant relatives with a risk tenfold lower in second-degree relatives and a risk comparable to the general population in third-degree relatives. On the basis of the pedigrees of IBD patients, segregation analyses were performed in the 1990s and proposed a major gene with dominant or recessive modes of inheritance in UC and CD respectively (Kuster *et al.*, 1989; Monsen, 1990; Orholm *et al.*, 1993).

The entire spectrum of IBD phenotypes may be encountered in familial cases. The phenotypic resemblance between affected relatives has been studied by many groups and most of them (but not all) retained a good concordance rate for age at onset and location of the disease between CD family members (Bayless *et al.*, 1996; Colombel *et al.*, 1996; Lee and Lennard-Jones, 1996; Satsangi *et al.*, 1996*a*). This result may suggest that IBD consists of a group of several similar phenotypes under the control of several Mendelian genes. In fact, such an interpretation has to be modified by the fact that extreme phenotypes such as CD and UC are often encountered in the same families. In addition, limited reports on the offspring of couples with two IBD parents suggest a genetic link between UC and CD (Bennett *et al.*, 1991; Laharie *et al.*, 2001). Indeed, more than 50% of the offspring of these couples develop an IBD before their twentieth birthday (Bennett *et al.*, 1991). Among them, the majority have a CD phenotype, even in the offspring of (CD/UC) and (UC/UC) couples. As a result, the rough concordance between affected relatives suggests some degree of genetic heterogeneity, which can also be seen at the phenotype level, but it does not allow a conclusion in favor of a patchwork of separate Mendelian disorders.

Genetic anticipation, the tendency for earlier presentation in successive generations, has been discussed in IBD following the initial report by Polito *et al.* (1996) that the age at onset (but not the severity) of the disease decreases by an average of 15 years with successive generations (Polito *et al.*, 1996). This difference has been widely confirmed but the explanation underlying the observation is the subject of debate. In fact, follow-up time biases and/or secular trends in disease presentation have been put forward as explanations rather than genetic anticipation.

Many individual cases of concordant twins and even concordant monozygotic triplets have been reported (Purrmann *et al.*, 1986). Twin studies performed on the basis of national registers are available from Sweden and Denmark ((Halfvarson *et al.*, 2003; Orholm *et al.*, 2000; Tysk *et al.*, 1988) and Table 20.2). The concordance rate between monozygotic twins was estimated to be 50% in CD and 16% in UC arguing for an environmental

Table 20.2. Summary of the twin studies from Sweden and Denmark

	Crohn's disease		Ulcerative colitis		Total cases
	Monozygotes	Dizygotes	Monozygotes	Dizygotes	
Sweden	9/18	1/26	3/16	0/20	80
Denmark	5/10	0/27	3/21	3/45	103

(Tysk *et al.*, 1988; Halfvarson *et al.*, 2003; Orholm *et al.*, 2000).

component for both diseases with a stronger effect in UC than in CD. In contrast, the concordance rate between dizygotic twins was estimated to be 2% in CD and 5% in UC. As a result, by comparing the two categories of twins, a strong genetic component for CD and a weaker one for UC can be proposed. Such a conclusion is in accordance with the lower percentage of familial aggregation seen in UC (see above).

Numerous genetically engineered mice have been proposed as human IBD models (for review see Bouma and Strober, 2003 and Table 20.3). In fact, mouse knockouts for genes as varied as IL10, Gαi2, NF-κB or N-cadherin lead to colitis. However, all of these models are not expected to be relevant to IBD and it appears more likely that colitis is a common phenotype in numerous disorders of the immune system or intestinal barrier. The mouse models most similar to CD are probably the SAMP1/Yit model and the mouse deleted for the Alu sequences of the TNFα promoter ($TNF^{\Delta ARE}$) which both develop a spontaneous ileitis with granuloma formation (Matsumoto *et al.*, 1998; Pizarro *et al.*, 2003). Interestingly, numerous animal models have some common features with IBD, including modulation of the phenotype by the intestinal flora and the occurrence of adenomas in the inflammatory lesions. Unfortunately, it was not possible until recently to apply the findings provided by these animal models to the human disease.

On the bases of evidence from twin and familial studies, geneticists have tried to identify IBD susceptibility genes. Many studies focused on candidate genes, most of them being immune regulatory genes (see Hugot *et al.*, 1999 and Table 20.4). Several associations have been found, most often in case-control studies. Unfortunately the majority of these associations were not confirmed in subsequent studies, making it difficult to understand if the first reports were false positive statistical results or genetic heterogeneity. Finally, many of the proposed associations were found with anonymous polymorphisms with no established functional effect. As a result, the literature on the candidate genes in IBD appears sometimes confusing and difficult to summarise. To date, the most convincing associations were reported for polymorphisms located in the major histocompatibility complex (where a linkage peak has also been demonstrated) and the Multidrug Resistance 1 gene (Brant *et al.*, 2003; Yap *et al.*, 2004).

Table 20.3. Genetically defined animal models developing a colitis

Knock-out mice	Transgenic animals	Spontaneous colitis
IL10	HLA B27	SAMP1/Yit
Mdr1a	Tgε26	C3H/HeJ/Bir
N-cadherin	IL7	
Trefoil peptide	$TNF^{\Delta ARE}$	
IL2/IL2Ra	Stat4	
CRFB4		
TGFβ/TGFβRII		
Gαi2		
TCRα		
NF-κB		
MHC Class II		
WASP		

Table 20.4. Functional candidate genes investigated in IBD

Candidate genes with an association reported at least once	Investigated genes with no reported association
HLA Class II	IL2
TNFα	PPARg
IL1/IL1ra	Complement
ICAM-1	TAP1/TAP2
Vitamin D receptor	Motilin
CCR5	TCR
CYP2D6	hMLH1
MDR1	Mucins
TNFR1/TNFR2	Interferon γ
	IL10
	IL4

Positional cloning of IBD genes

Using a positional cloning approach, several investigators were able to detect linkage to IBD on different chromosomes (Cho *et al.*, 1998; Duerr *et al.*, 2000; Hampe *et al.*, 1999*a*; Hampe *et al.*, 1999*b*; Hugot *et al.*, 1996; Rioux *et al.*, 2000; Satsangi *et al.*, 1996*b*). Nine IBD loci are today accepted as firmly established according to the criteria defined by Kruglyak and Lander (Figure 20.1). Additional loci exhibit suggestive evidence of linkage. Finally, a meta-analysis of the published genome scans revealed a significant linkage in the HLA region and suggestive linkage on the previously reported regions 5q, 7q and 16 and on three additional regions: 2q (IBD), 3q (CD) and 17q (CD) (van Heel *et al.*, 2004). Among all the reported loci, not one has an attributable λs value higher than 2, indicating that IBD are complex genetic traits with no major gene. Most of the loci (all but *IBD2*) seem to be involved in CD. In contrast, several loci do not seem to play a significant role in UC (*IBD1, 4, 5* and *DGL5*), arguing again in favour of a stronger genetic component for CD than for UC. Finally, among the nine IBD loci, three genes involved in CD predisposition have been identified using a positional cloning approach: *CARD15*, *SLC22A4/5* and *DLG5* (see below).

On chromosome 5, the *IBD5* locus was localized to a region of 250 kb within an area containing several cytokine genes (Rioux *et al.*, 2000; 2001). In fact, the identified genes (*SLC22A4* and *SLC22A5*) do not code for cytokines but for the organic cation transporters OCTN1 and OCTN2 respectively (Peltekova *et al.*, 2004). Two genetic variants in strong linkage disequilibrium with each other have been associated with CD. The haplotype defined by these two variants is present in 42% of healthy controls and 54% of CD patients. The associated protein seems to interact with *CARD15/NOD2*. The first genetic variant codes for a non-conservative amino acid substitution (L503F) in the OCTN1 protein. This variant alters the Vmax and Km constants of the transporter toward several substrates, including carnitine and xenobiotics. The second genetic variant (*-207G>C*) affects a heat shock transcription factor element in the promoter of *SLC22A5* resulting in decreased transcription of the gene when activated. The two genes are widely expressed but their function in the context of IBD remains highly speculative. However, it should be noted that a genetic polymorphism located in a promoter RUNX1 element of *SLC22A4* has been associated with rheumatoid arthritis, providing a link between these two inflammatory disorders (Tokuhiro *et al.*, 2003). As a result, the OCTN genes may have an unexpected but important role in immune disorders.

A gene on chromosomal band 10q23 has also been identified recently (Stoll *et al.*, 2004). Two main variants within the *DLG5* gene coding for non-conservative substitutions of the protein (R30Q and P1371Q) have been associated with CD and possibly with UC. The first mutation is relatively frequent in the general population (34%) while the second one is rare (less than 1%). This observation was made using case-control and transmission disequilibrium analyses in several independent White populations, together with reports of additional private variants. An interaction with *CARD15/NOD2* is likely. *DLG5* encodes a

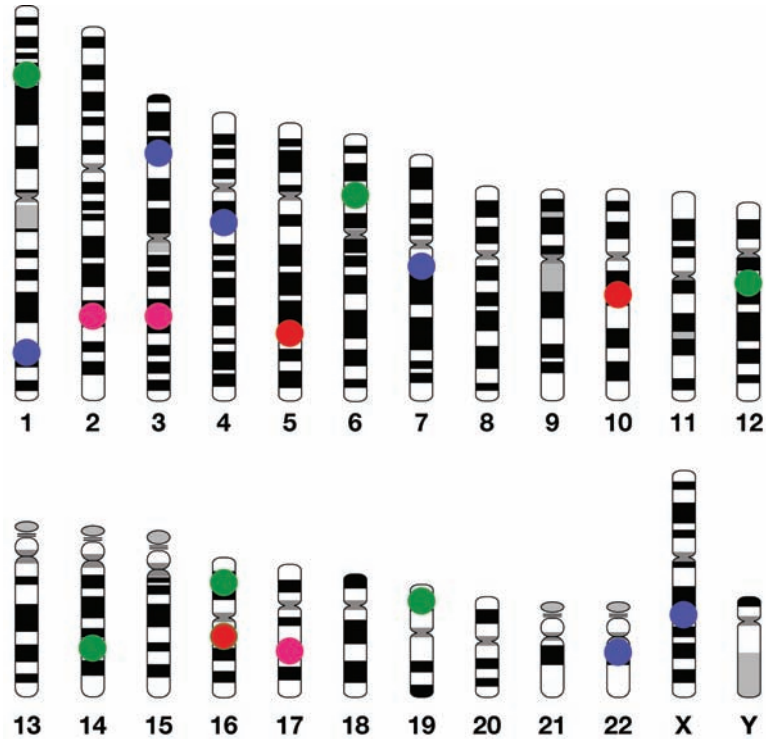

Figure 20.1 Localization of the *IBD* loci. Three genes (in red) playing a role in CD have been identified by a positional cloning approach: *CARD15/NOD2, SLC22A4/5* and *DLG5. 6 loci* (in green) have shown significant linkage while other loci have only demonstrated suggestive linkage in single genome-scan analyses (in blue) or in the meta-analysis of the reported genome-scans (in pink).

scaffolding protein that may be important for the maintenance of epithelial integrity. It contains several PDZ domains. Here too, the functional link between this gene and the IBD phenotype remains to be explored. In fact, the discoveries of *DLG5* and the genes coding for the OCTN transporters demonstrate the power of positional cloning approach in IBD: these genes would certainly not have been investigated as candidate genes on the basis of their known functions.

A major breakthrough: identification of the *CARD15/NOD2* gene

The best known of the IBD genes is *CARD15/NOD2*. The association between this gene and CD was provided simultaneously by two groups and further confirmed by many additional investigators (Hugot *et al.*, 2001; Ogura *et al.*, 2001*a*). The gene is located on chromosome band 16q12 and corresponds to the *IBD1* locus. About 60 polymorphisms have been reported within the gene including 30 mutations coding for non-conservative amino acid substitutions in the protein (Lesage *et al.*, 2002). However, most *CARD15/NOD2* mutations are private and more than 80% of mutated chromosomes include three main variants: R702W, G908R and 1007fs. In more than 300 IBD families, we failed to detect a chromosome carrying two of these variants (Hugot *et al.*, 2001). We were thus able to demonstrate that the three major mutations are independently associated with the disease.

Unexpectedly, all three variants occurred on the same ancestral haplotype, which is not by itself associated with CD (Hugot *et al.*, 2001). Finally, we were able to demonstrate that families exhibiting either the common or private mutations explained the entire linkage at the *IBD1* locus. As a whole, these genetic analyses provided strong evidence that *CARD15/NOD2* is a CD susceptibility gene.

The rather simple mutational spectrum of the gene provided the opportunity to easily build genetic testing procedures for CD diagnostics. As expected for a complex genetic disorder, *CARD15/NOD2* was found to be neither necessary nor sufficient for disease development. Considering the private and frequent mutations together, around 50% of CD patients and 20% of healthy controls were found to carry one or more mutations (Lesage *et al.*, 2002). UC patients are not different from healthy controls. Interestingly, the percentage of CD patients (17%) carrying two mutations (on their two chromosomes 16) was found to be much higher than expected by chance alone (Lesage *et al.*, 2002). This mutation dosage effect is reminiscent of the recessive model proposed by segregation analyses for CD. It also suggested a loss-of-function model for CD associated with *CARD15/NOD2* mutations.

Despite these observations, genetic testing is not at present proposed for routine diagnosis. First of all, there is no preventive action that can be proposed for genetically at-risk individuals. As a result, a pre-symptomatic diagnostic procedure is of doubtful value. Moreover, the odds ratio for disease occurrence is in the range of 2 to 3 for single mutation carriers and 20 to 40 for double mutation carriers (Hampe *et al.*, 2001; Hugot *et al.*, 2001; Ogura *et al.*, 2001*a*). Considering a rough CD prevalence of 1/1000, it can be deduced that the increased risk attributable to these *CARD15/NOD2* genotypes is in the range of a few percent. Such information is too limited to be useful in practice.

In order to define the relevance of genotype testing during a diagnostic procedure, the genotype/phenotype relationship was studied by many groups (Abreu *et al.*, 2002; Ahmad *et al.*, 2002; Cuthbert *et al.*, 2002; Hampe *et al.*, 2002; Helio *et al.*, 2003; Lesage *et al.*, 2002; Louis *et al.*, 2003; Murillo *et al.*, 2002; Radlmayr *et al.*, 2002; Vermeire *et al.*, 2002*b*). Mutation carriers (and especially double mutation carriers) are usually characterized by a younger age at onset, by ileal involvement and by more frequent local complications such as stenoses and fistulas. Unfortunately, this information is not always helpful in clinical practice. The impact of genetic testing is also limited in terms of therapeutic choice (Mascheretti *et al.*, 2002; Vermeire *et al.*, 2002*a*). As a result, genetic testing is not currently proposed in routine clinical practice. However, this position will certainly be revisited in the near future due to the recent discovery of additional CD susceptibility genes, such as the *DLG5* and *OCTN* genes. Indeed, we hope that a molecular diagnostics kit will be available in the next few years.

The CARD15/NOD2 protein contains three different parts: a leucine-rich repeat domain (LRR) with a horseshoe structure; a nucleotide binding domain (NBD) involved in protein self-oligomerization and two caspase recruitment domains (CARDs) usually involved in homophilic CARD/CARD interactions (Ogura *et al.*, 2001*b*). *CARD15/NOD2* therefore has a structure very similar to proteins involved in innate immunity, including plant proteins, such as that involved in resistance to tobacco mosaic virus. As expected for a gene involved in innate immunity, *CARD15/NOD2* is mainly expressed in white blood cells, including monocytes, macrophages and dendritic cells (Gutierrez *et al.*, 2002). *CARD15/NOD2* can also be expressed by enterocytes after activation by inflammatory cytokines, such as TNFα or IFNγ (Gutierrez *et al.*, 2002; Rosenstiel *et al.*, 2003). Among the enterocytes, specialized cells, known as Paneth cells seem to express *CARD15/NOD2* at a high level (Lala *et al.*, 2003; Ogura *et al.*, 2003). These cells are involved in local immunity, secreting lysozyme and defensins. As a result, *CARD15/NOD2* is a good candidate to explain the proposed link between the host and intestinal flora.

CARD15/NOD2 was initially thought to drive the response of the host against bacterial lipopolysaccharide (LPS) (Inohara *et al.*, 2001; Ogura *et al.*, 2001*b*). Further experiments demonstrated that *CARD15/NOD2* is a sensor for the peptidoglycan (PG) present in the bacterial cell wall (Girardin *et al.*, 2003*a*; 2003*b*; Inohara *et al.*, 2003). The minimal structures recognized by *CARD15/NOD2* are muropeptides including muramyl dipeptide (MDP) which has been used for a long time as an adjuvant for vaccinations. However, despite the clear demonstration of the activation of *CARD15/NOD2* by the MDP, several questions remain. First, *CARD15/NOD2* is present in the cytoplasm of cells. Considering that the muropeptides are hydrophilic components they are not expected to cross the cell membrane and little is known about how MDP enters the cell. Second, recent in silico analyses suggest that MDP is too small to interact alone with the LRR of *CARD15/NOD2*. As a result, additional protein(s) are expected to play a role in the *CARD15/NOD2*–MDP interaction. Under activation by the muropeptides, *CARD15/NOD2* is able to activate the NF-κB pathway (Inohara *et al.*, 2001; Ogura *et al.*, 2001*b*). This effect seems to be due to the self-oligomerization of the NOD domains followed by the recruitment of the Rick/Rip2 protein by homophilic CARD/CARD interactions (Chamaillard *et al.*, 2003; Inohara *et al.*, 2000). The result of this NF-κB activation is expected to have a pro-inflammatory effect through the transcription of various cytokines, including TNFα and IL10. This provides a link between the gut flora and the inflammatory disorder.

Functional studies were carried out to try and confirm this proposal. Most of the available data have been provided by relatively crude models of NF-κB activation in transiently transfected cell lines. They have therefore to be interpreted with caution. Despite these limitations, it is widely accepted that most (but not all) the CD-associated mutations are characterized by a loss of NF-κB activation in the presence of MDP (Kobayashi *et al.*, 2002; Ogura *et al.*, 2001*a*). Consequently, it looks as if the loss of the pro-inflammatory pathway defined by *CARD15/NOD2* results in CD, which is an inflammatory disorder.

This paradox is not fully explained today. Most researchers think that this loss of function induces a shift towards an adaptive immune response, resulting in a more intense reaction and a loss of tolerance in the gut. However, this hypothesis has not yet been confirmed. Another explanation could be that the *CARD15/NOD2* pathway also has anti-inflammatory properties. Recent reports showed that *CARD15/NOD2* interacts with alternative pathways able to activate NF-κB. For example, *CARD15/NOD2* is able to inhibit *TAK1* (Chen *et al.*, 2004). It is also able to modulate the Toll-like Receptor 2 (*TLR2*) response to PG, resulting in inhibition of the expression of IFNγ and IL12 (Netea *et al.*, 2004; Watanabe *et al.*, 2004). These recent data argue for studies under more physiological conditions before coming to firm conclusions (Hisamatsu *et al.*, 2003; Opitz *et al.*, 2004). Finally, the loss-of-function model of *CARD15/NOD2* can also be questioned. All CD-associated mutations do not have the same functional defect, suggesting that the experiments performed to date are not fully relevant (Chamaillard *et al.*, 2003). In addition, *CARD15/NOD2* knockout mice do not exhibit a definite phenotype. Specifically, they do not develop any digestive disease (Pauleau and Murray, 2003).

Interestingly, genetic variation in *CARD15/NOD2* is not only associated with CD but also with additional disorders. CD-associated mutations have been proposed to have a role in asthma and more recently in graft versus host disease (Holler *et al.*, 2004). These observations show that a gene involved in a complex genetic disorder can also be important for other disorders. As a result, it is important to define the entire spectrum of phenotypes associated with a single gene. In this respect, no association was found between *CARD15/NOD2* and spondyloarthropathies, rheumatoid arthritis or Wegener disease.

CARD15/NOD2 is also involved in another disease characterized by inflammation with granuloma formation: Blau syndrome (BS) (Miceli-Richard *et al.*, 2001). BS is a rare autosomal

dominant Mendelian disorder defined by relapsing inflammation, including skin rashes, uveitis and arthritis. There is no reported digestive lesion in BS. We found that BS is characterized by mutations affecting the NBD region of the protein, especially at codons 334 and 469 (Miceli-Richard *et al.*, 2001). At a functional level, these mutations are characterized by a gain-of-function: the mutated protein is able to activate NF-κB without any stimulation by the MDP (Chamaillard *et al.*, 2003). In contrast to CD, the findings in BS are thus consistent with the observed auto-inflammatory phenotype. Considering that MDP is not required for activation of the pathway, the involvement of sterile intestinal sites is also consistent with a gain-of-function model. Interestingly, BS mutations are very similar to those occurring in the *PYPAF1* gene, which is responsible for the Muckle–Wells syndrome (MWS), familial cold autoinflammatory syndrome (FCAS) and chronic infantile neurological cutaneous and articular syndrome (CINCA) (Albrecht *et al.*, 2003).

It is tempting to try and reconcile the findings observed in CD and BS. In this respect, we recently proposed a model of *CARD15/NOD2* inhibition by bacterial components (Hugot *et al.*, 2003*a*). In fact, IBD perfectly illustrates the difficulties inherent in studies of complex genetic disorders characterized by the interaction of many variables. Although we were able to identify three IBD susceptibility genes using a positional cloning approach, there is still a long way to go before we can fill the gap between the genetic data and a relevant model explaining the disease mechanisms.

REFERENCES

Abreu, M. T., Taylor, K. D., Lin, Y. C. *et al.* (2002). Mutations in NOD2 are associated with fibrostenosing disease in patients with Crohn's disease. *Gastroenterology*, **123**, 679–88.

Ahmad, T., Armuzzi, A., Bunce, M. *et al.* (2002). The molecular classification of the clinical manifestations of Crohn's disease. *Gastroenterology*, **122**, 854–66.

Albrecht, M., Domingues, F. S., Schreiber, S. and Lengauer, T. (2003). Structural localization of disease-associated sequence variations in the NACHT and LRR domains of PYPAF1 and NOD2. *FEBS Lett*, **554**, 520–8.

Andersson, R. E., Olaison, G., Tysk, C. and Ekbom, A. (2001). Appendectomy and protection against ulcerative colitis. *N Engl J Med*, **344**, 808–14.

Bayless, T. M., Tokayer, A. Z., Polito, J. M., 2nd, *et al.* (1996). Crohn's disease: concordance for site and clinical type in affected family members–potential hereditary influences. *Gastroenterology*, **111**, 573–9.

Bennett, R. A., Rubin, P. H. and Present, D. H. (1991). Frequency of inflammatory bowel disease in offspring of couples both presenting with inflammatory bowel disease. *Gastroenterology*, **100**, 1638–43.

Bouma, G. and Strober, W. (2003). The immunological and genetic basis of inflammatory bowel disease. *Nat Rev Immunol*, **3**, 521–33.

Brahme, F., Lindstrom, C. and Wenckert, A. (1975). Crohn's disease in a defined population. An epidemiological study of incidence, prevalence, mortality, and secular trends in the city of Malmo, Sweden. *Gastroenterology*, **69**, 342–51.

Brant, S. R., Panhuysen, C. I., Nicolae, D. *et al.* (2003). MDR1 Ala893 polymorphism is associated with inflammatory bowel disease. *Am J Hum Genet*, **73**, 1282–92.

Bridger, S., Lee, J. C., Bjarnason, I., Jones, J. E. and Macpherson, A. J. (2002). In siblings with similar genetic susceptibility for inflammatory bowel disease, smokers tend to develop Crohn's disease and non-smokers develop ulcerative colitis. *Gut*, **51**, 21–5.

Bull, T. J., McMinn, E. J., Sidi-Boumedine, K. *et al.* (2003). Detection and verification of *Mycobacterium avium* subsp. paratuberculosis in fresh ileocolonic mucosal biopsy specimens from individuals with and without Crohn's disease. *J Clin Microbiol*, **41**, 2915–23.

Calkins, B. M. (1989). A meta-analysis of the role of smoking in inflammatory bowel disease. *Dig Dis Sci*, **34**, 1841–54.

Chamaillard, M., Philpott, D., Girardin, S. E. *et al.* (2003). Gene-environment interaction modulated by allelic heterogeneity in inflammatory diseases. *Proc Natl Acad Sci USA*, **100**, 3455–60.

Chen, C. M., Gong, Y., Zhang, M. and Chen, J. J. (2004). Reciprocal cross-talk between *NOD2* and *TAK1* signaling pathways. *J Biol Chem*, **279**, 25876–82.

Cho, J. H., Nicolae, D. L., Gold, L. H. *et al.* (1998). Identification of novel susceptibility loci for inflammatory

bowel disease on chromosomes 1p, 3q, and 4q: evidence for epistasis between 1p and IBD1. *Proc Natl Acad Sci USA*, **95**, 7502–7.

Colombel, J. F., Grandbastien, B., Gower-Rousseau, C. *et al.* (1996). Clinical characteristics of Crohn's disease in 72 families. *Gastroenterology*, **111**, 604–7.

Cosnes, J., Beaugerie, L., Carbonnel, F. and Gendre, J. P. (2001). Smoking cessation and the course of Crohn's disease: an intervention study. *Gastroenterology*, **120**, 1093–9.

Cosnes, J., Carbonnel, F., Beaugerie, L., Le Quintrec, Y. and Gendre, J. P. (1996). Effects of cigarette smoking on the long-term course of Crohn's disease. *Gastroenterology*, **110**, 424–31.

Cosnes, J., Carbonnel, F., Carrat, F., Beaugerie, L. and Gendre, J. P. (1999). Oral contraceptive use and the clinical course of Crohn's disease: a prospective cohort study. *Gut*, **45**, 218–22.

Cottone, M., Rosselli, M., Orlando, A. *et al.* (1994). Smoking habits and recurrence in Crohn's disease. *Gastroenterology*, **106**, 643–8.

Cuthbert, A. P., Fisher, S. A., Mirza, M. M. *et al.* (2002). The contribution of *NOD2* gene mutations to the risk and site of disease in inflammatory bowel disease. *Gastroenterology*, **122**, 867–74.

Darfeuille-Michaud, A., Neut, C., Barnich, N. *et al.* (1998). Presence of adherent *Escherichia coli* strains in ileal mucosa of patients with Crohn's disease. *Gastroenterology*, **115**, 1405–13.

Duerr, R. H., Barmada, M. M., Zhang, L., Pfutzer, R. and Weeks, D. E. (2000). High-density genome scan in Crohn disease shows confirmed linkage to chromosome 14q11–12. *Am J Hum Genet*, **66**, 1857–62.

Ekbom, A., Adami, H. O., Helmick, C. G., Jonzon, A. and Zack, M. M. (1990). Perinatal risk factors for inflammatory bowel disease: a case-control study. *Am J Epidemiol*, **132**, 1111–19.

Ekbom, A., Daszak, P., Kraaz, W. and Wakefield, A. J. (1996). Crohn's disease after in-utero measles virus exposure. *Lancet*, **348**, 515–17.

Evans, J. G. and Acheson, E. D. (1965). An epidemiological study of ulcerative colitis and regional enteritis in the Oxford area. *Gut*, **6**, 311–24.

Fellows, I. W., Freeman, J. G. and Holmes, G. K. (1990). Crohn's disease in the city of Derby, 1951–85. *Gut*, **31**, 1262–5.

Gent, A. E., Hellier, M. D., Grace, R. H., Swarbrick, E. T. and Coggon, D. (1994). Inflammatory bowel disease and domestic hygiene in infancy. *Lancet*, **343**, 766–7.

Gilat, T., Hacohen, D., Lilos, P. and Langman, M. J. (1987). Childhood factors in ulcerative colitis and Crohn's disease. An international cooperative study. *Scand J Gastroenterol*, **22**, 1009–24.

Girardin, S. E., Boneca, I. G., Viala, J. *et al.* (2003*a*). NOD2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection. *J Biol Chem*, **278**, 8869–72.

Girardin, S. E., Travassos, L. H., Herve, M. *et al.* (2003*b*). Peptidoglycan molecular requirements allowing detection by *Nod1* and *NOD2*. *J Biol Chem*, **278**, 41702–8.

Godet, P. G., May, G. R. and Sutherland, L. R. (1995). Meta-analysis of the role of oral contraceptive agents in inflammatory bowel disease. *Gut*, **37**, 668–73.

Greenstein, R. J. (2003). Is Crohn's disease caused by a mycobacterium? Comparisons with leprosy, tuberculosis, and Johne's disease. *Lancet Infect Dis*, **3**, 507–14.

Gutierrez, O., Pipaon, C., Inohara, N. *et al.* (2002). Induction of *NOD2* in myelomonocytic and intestinal epithelial cells via nuclear factor-kappa B activation. *J Biol Chem*, **277**, 41701–5.

Halfvarson, J., Bodin, L., Tysk, C., Lindberg, E. and Jarnerot, G. (2003). Inflammatory bowel disease in a Swedish twin cohort: a long-term follow-up of concordance and clinical characteristics. *Gastroenterology*, **124**, 1767–73.

Hampe, J., Cuthbert, A., Croucher, P. J. *et al.* (2001). Association between insertion mutation in *NOD2* gene and Crohn's disease in German and British populations. *Lancet*, **357**, 1925–8.

Hampe, J., Grebe, J., Nikolaus, S. *et al.* (2002). Association of *NOD2* (*CARD 15*) genotype with clinical course of Crohn's disease: a cohort study. *Lancet*, **359**, 1661–5.

Hampe, J., Schreiber, S., Shaw, S. H. *et al.* (1999*a*). A genomewide analysis provides evidence for novel linkages in inflammatory bowel disease in a large European cohort. *Am J Hum Genet*, **64**, 808–16.

Hampe, J., Shaw, S. H., Saiz, R. *et al.* (1999*b*). Linkage of inflammatory bowel disease to human chromosome 6p. *Am J Hum Genet*, **65**, 1647–55.

Helio, T., Halme, L., Lappalainen, M. *et al.* (2003). *CARD15/NOD2* gene variants are associated with familially occurring and complicated forms of Crohn's disease. *Gut*, **52**, 558–62.

Hellers, G. (1979). Crohn's disease in Stockholm county 1955–1974. A study of epidemiology, results of surgical treatment and long-term prognosis. *Acta Chir Scand Suppl*, **490**, 1–84.

Hisamatsu, T., Suzuki, M., Reinecker, H. C. *et al.* (2003). *CARD15/NOD2* functions as an antibacterial factor in human intestinal epithelial cells. *Gastroenterology*, **124**, 993–1000.

Holler, E., Rogler, G., Herfarth, H. *et al.* (2004). Both donor and recipient *NOD2/CARD15* mutations associate with transplant-related mortality and GvHD following allogeneic stem cell transplantation. *Blood*, **104**, 889–94.

Hugot, J. P., Alberti, C., Berrebi, D., Bingen, E. and Cezard, J. P. (2003*a*). Crohn's disease: the cold chain hypothesis. *Lancet*, **362**, 2012–15.

Hugot, J. P., Cezard, J. P., Colombel, J. F. *et al.* (2003*b*). Clustering of Crohn's disease within affected sibships. *Eur J Hum Genet*, **11**, 179–84.

Hugot, J. P., Chamaillard, M., Zouali, H. *et al.* (2001). Association of *NOD2* leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature*, **411**, 599–603.

Hugot, J. P., Laurent-Puig, P., Gower-Rousseau, C. *et al.* (1996). Mapping of a susceptibility locus for Crohn's disease on chromosome 16. *Nature*, **379**, 821–3.

Hugot, J. P., Zouali, H., Lesage, S. and Thomas, G. (1999). Etiology of the inflammatory bowel diseases. *Int J Colorectal Dis*, **14**, 2–9.

Inohara, N., Koseki, T., Lin, J. *et al.* (2000). An induced proximity model for NF-kappa B activation in the *Nod1/RICK* and *RIP* signaling pathways. *J Biol Chem*, **275**, 27823–31.

Inohara, N., Ogura, Y., Chen, F. F., Muto, A. and Nunez, G. (2001). Human *Nod1* confers responsiveness to bacterial lipopolysaccharides. *J Biol Chem*, **276**, 2551–4.

Inohara, N., Ogura, Y., Fontalba, A. *et al.* (2003). Host recognition of bacterial muramyl dipeptide mediated through *NOD2*. Implications for Crohn's disease. *J Biol Chem*, **278**, 5509–12.

Kallinowski, F., Wassmer, A., Hofmann, M. A. *et al.* (1998). Prevalence of enteropathogenic bacteria in surgically treated chronic inflammatory bowel disease. *Hepatogastroenterology*, **45**, 1552–8.

Kobayashi, K., Inohara, N., Hernandez, L. D. *et al.* (2002). RICK/Rip2/CARDIAK mediates signalling for receptors of the innate and adaptive immune systems. *Nature*, **416**, 194–9.

Kugathasan, S., Judd, R. H., Hoffmann, R. G. *et al.* (2003). Epidemiologic and clinical characteristics of children with newly diagnosed inflammatory bowel disease in Wisconsin: a statewide population-based study. *J Pediatr*, **143**, 525–31.

Kuster, W., Pascoe, L., Purrmann, J., Funk, S. and Majewski, F. (1989). The genetics of Crohn disease: complex segregation analysis of a family study with 265 patients with Crohn disease and 5,387 relatives. *Am J Med Genet*, **32**, 105–8.

Kyle, J. (1971). An epidemiological study of Crohn's disease in Northeast Scotland. *Gastroenterology*, **61**, 826–33.

Laharie, D., Debeugny, S., Peeters, M. *et al.* (2001). Inflammatory bowel disease in spouses and their offspring. *Gastroenterology*, **120**, 816–19.

Lala, S., Ogura, Y., Osborne, C. *et al.* (2003). Crohn's disease and the *NOD2* gene: a role for paneth cells. *Gastroenterology*, **125**, 47–57.

Lamps, L. W., Madhusudhan, K. T., Havens, J. M. *et al.* (2003). Pathogenic *Yersinia* DNA is detected in bowel and mesenteric lymph nodes from patients with Crohn's disease. *Am J Surg Pathol*, **27**, 220–7.

Lee, J. C. and Lennard-Jones, J. E. (1996). Inflammatory bowel disease in 67 families each with three or more affected first-degree relatives. *Gastroenterology*, **111**, 587–96.

Lesage, S., Zouali, H., Cezard, J. P. *et al.* (2002). *CARD15/NOD2* mutational analysis and genotype–phenotype correlation in 612 patients with inflammatory bowel disease. *Am J Hum Genet*, **70**, 845–57.

Lindberg, E., Jarnerot, G. and Huitfeldt, B. (1992). Smoking in Crohn's disease: effect on localisation and clinical course. *Gut*, **33**, 779–82.

Loftus, E. V., Jr., Silverstein, M. D., Sandborn, W. J. *et al.* (1998). Crohn's disease in Olmsted County, Minnesota, 1940–1993: incidence, prevalence, and survival. *Gastroenterology*, **114**, 1161–8.

Louis, E., Michel, V., Hugot, J. P. *et al.* (2003). Early development of stricturing or penetrating pattern in Crohn's disease is influenced by disease location, number of flares, and smoking but not by *NOD2/CARD15* genotype. *Gut*, **52**, 552–7.

Mascheretti, S., Hampe, J., Croucher, P. J. *et al.* (2002). Response to infliximab treatment in Crohn's disease is not associated with mutations in the *CARD15 (NOD2)* gene: an analysis in 534 patients from two multicenter, prospective GCP-level trials. *Pharmacogenetics*, **12**, 509–15.

Matsumoto, S., Okabe, Y., Setoyama, H. *et al.* (1998). Inflammatory bowel disease-like enteritis and caecitis in a senescence accelerated mouse P1/Yit strain. *Gut*, **43**, 71–8.

Mayberry, J.F. and Rhodes, J. (1984). Epidemiological aspects of Crohn's disease: a review of the literature. *Gut*, **25**, 886–99.

Miceli-Richard, C., Lesage, S., Rybojad, M. *et al.* (2001). *CARD15* mutations in Blau syndrome. *Nat Genet*, **29**, 19–20.

Monsen, U. (1990). Inflammatory bowel disease. An epidemiological and genetic study. *Acta Chir Scand Suppl*, **559**, 1–42.

Monsen, U., Brostrom, O., Nordenvall, B., Sorstad, J. and Hellers, G. (1987). Prevalence of inflammatory bowel disease among relatives of patients with ulcerative colitis. *Scand J Gastroenterol*, **22**, 214–18.

Murillo, L., Crusius, J.B., van Bodegraven, A.A., Alizadeh, B.Z. and Pena, A.S. (2002). *CARD15* gene and the classification of Crohn's disease. *Immunogenetics*, **54**, 59–61.

Netea, M.G., Kullberg, B.J., de Jong, D.J. *et al.* (2004). *NOD2* mediates anti-inflammatory signals induced by TLR2 ligands: implications for Crohn's disease. *Eur J Immunol*, **34**, 2052–9.

Odes, H.S., Fich, A., Reif, S. *et al.* (2001). Effects of current cigarette smoking on clinical course of Crohn's disease and ulcerative colitis. *Dig Dis Sci*, **46**, 1717–21.

Ogura, Y., Bonen, D.K., Inohara, N. *et al.* (2001*a*). A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature*, **411**, 603–6.

Ogura, Y., Inohara, N., Benito, A. *et al.* (2001*b*). NOD2, a *Nod1/Apaf-1* family member that is restricted to monocytes and activates NF-kappaB. *J Biol Chem*, **276**, 4812–18.

Ogura, Y., Lala, S., Xin, W. *et al.* (2003). Expression of *NOD2* in Paneth cells: a possible link to Crohn's ileitis. *Gut*, **52**, 1591–7.

Opitz, B., Puschel, A., Schmeck, B. *et al.* (2004). Nucleotide-binding oligomerization domain proteins are innate immune receptors for internalized *Streptococcus pneumoniae*. *J Biol Chem*, **279**, 36426–32.

Orholm, M., Binder, V., Sorensen, T.I., Rasmussen, L.P. and Kyvik, K.O. (2000). Concordance of inflammatory bowel disease among Danish twins. Results of a nationwide study. *Scand J Gastroenterol*, **35**, 1075–81.

Orholm, M., Iselius, L., Sorensen, T.I. *et al.* (1993). Investigation of inheritance of chronic inflammatory bowel diseases by complex segregation analysis. *British Medical Journal*, **306**, 20–4.

Pauleau, A.L. and Murray, P.J. (2003). Role of *NOD2* in the response of macrophages to Toll-like receptor agonists. *Mol Cell Biol*, **23**, 7531–9.

Peeters, M., Nevens, H., Baert, F. *et al.* (1996). Familial aggregation in Crohn's disease: increased age-adjusted risk and concordance in clinical characteristics. *Gastroenterology*, **111**, 597–603.

Peltekova, V.D., Wintle, R.F., Rubin, L.A. *et al.* (2004). Functional variants of OCTN cation transporter genes are associated with Crohn disease. *Nat Genet*, **36**, 471–5.

Persson, P.G., Ahlbom, A. and Hellers, G. (1987). Crohn's disease and ulcerative colitis. A review of dietary studies with emphasis on methodologic aspects. *Scand J Gastroenterol*, **22**, 385–9.

Pizarro, T.T., Arseneau, K.O., Bamias, G. and Cominelli, F. (2003). Mouse models for the study of Crohn's disease. *Trends Mol Med*, **9**, 218–22.

Podolsky, D.K. (2002). Inflammatory bowel disease. *N Engl J Med*, **347**, 417–29.

Polito, J.M., 2nd, Rees, R.C., Childs, B. *et al.* (1996). Preliminary evidence for genetic anticipation in Crohn's disease. *Lancet*, **347**, 798–800.

Probert, C.S., Jayanthi, V., Pinder, D., Wicks, A.C. and Mayberry, J.F. (1992). Epidemiological study of ulcerative proctocolitis in Indian migrants and the indigenous population of Leicestershire. *Gut*, **33**, 687–93.

Purrmann, J., Bertrams, J., Borchard, F. *et al.* (1986). Monozygotic triplets with Crohn's disease of the colon. *Gastroenterology*, **91**, 1553–9.

Radlmayr, M., Torok, H.P., Martin, K. and Folwaczny, C. (2002). The c-insertion mutation of the *NOD2* gene is associated with fistulizing and fibrostenotic phenotypes in Crohn's disease. *Gastroenterology*, **122**, 2091–2.

Rioux, J.D., Daly, M.J., Silverberg, M.S. *et al.* (2001). Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. *Nat Genet*, **29**, 223–8.

Rioux, J.D., Silverberg, M.S., Daly, M.J. *et al.* (2000). Genomewide search in Canadian families with inflammatory bowel disease reveals two novel susceptibility loci. *Am J Hum Genet*, **66**, 1863–70.

Rose, J.D., Roberts, G.M., Williams, G., Mayberry, J.F. and Rhodes, J. (1988). Cardiff Crohn's disease jubilee: the incidence over 50 years. *Gut*, **29**, 346–51.

Rosenstiel, P., Fantini, M., Brautigam, K. *et al.* (2003). TNF-alpha and IFN-gamma regulate the expression of the *NOD2 (CARD15)* gene in human intestinal epithelial cells. *Gastroenterology*, **124**, 1001–9.

Roth, M. P., Petersen, G. M., McElree, C. *et al.* (1989). Familial empiric risk estimates of inflammatory bowel disease in Ashkenazi Jews. *Gastroenterology*, **96**, 1016–20.

Russell, R. K. and Satsangi, J. (2004). IBD: a family affair. *Best Pract Res Clin Gastroenterol*, **18**, 525–39.

Sanderson, J. D., Moss, M. T., Tizard, M. L. and Hermon-Taylor, J. (1992). *Mycobacterium paratuberculosis* DNA in Crohn's disease tissue. *Gut*, **33**, 890–6.

Satsangi, J., Grootscholten, C., Holt, H. and Jewell, D. P. (1996*a*). Clinical patterns of familial inflammatory bowel disease. *Gut*, **38**, 738–41.

Satsangi, J., Parkes, M., Louis, E. *et al.* (1996*b*). Two stage genome-wide search in inflammatory bowel disease provides evidence for susceptibility loci on chromosomes 3, 7 and 12. *Nat Genet*, **14**, 199–202.

Shanahan, F. (2002). Crohn's disease. *Lancet*, **359**, 62–9.

Shivananda, S., Lennard-Jones, J., Logan, R. *et al.* (1996). Incidence of inflammatory bowel disease across Europe: is there a difference between north and south? Results of the European Collaborative Study on Inflammatory Bowel Disease (EC-IBD). *Gut*, **39**, 690–7.

Stoll, M., Corneliussen, B., Costello, C. M. *et al.* (2004). Genetic variation in *DLG5* is associated with inflammatory bowel disease. *Nat Genet*, **36**, 476–80.

Sutherland, L. R., Ramcharan, S., Bryant, H. and Fick, G. (1990). Effect of cigarette smoking on recurrence of Crohn's disease. *Gastroenterology*, **98**, 1123–8.

Thompson, N. P., Montgomery, S. M., Pounder, R. E. and Wakefield, A. J. (1995). Is measles vaccination a risk factor for inflammatory bowel disease? *Lancet*, **345**, 1071–4.

Tokuhiro, S., Yamada, R., Chang, X. *et al.* (2003). An intronic SNP in a RUNX1 binding site of SLC22A4, encoding an organic cation transporter, is associated with rheumatoid arthritis. *Nat Genet*, **35**, 341–8.

Tysk, C., Lindberg, E., Jarnerot, G. and Floderus-Myrhed, B. (1988). Ulcerative colitis and Crohn's disease in an unselected population of monozygotic and dizygotic twins. A study of heritability and the influence of smoking. *Gut*, **29**, 990–6.

van Heel, D. A., Fisher, S. A., Kirby, A. *et al.* (2004). Inflammatory bowel disease susceptibility loci defined by genome scan meta-analysis of 1952 affected relative pairs. *Hum Mol Genet*, **13**, 763–70.

Vermeire, S., Louis, E., Rutgeerts, P. *et al.* (2002*a*). *NOD2/CARD15* does not influence response to infliximab in Crohn's disease. *Gastroenterology*, **123**, 106–11.

Vermeire, S., Wild, G., Kocher, K. *et al.* (2002*b*). *CARD15* genetic variation in a Quebec population: prevalence, genotype–phenotype relationship, and haplotype structure. *Am J Hum Genet*, **71**, 74–83.

Watanabe, T., Kitani, A., Murray, P. J. and Strober, W. (2004). *NOD2* is a negative regulator of Toll-like receptor 2-mediated T helper type 1 responses. *Nat Immunol*, **5**, 800–8.

Yap, L. M., Ahmad, T. and Jewell, D. P. (2004). The contribution of HLA genes to IBD susceptibility and phenotype. *Best Pract Res Clin Gastroenterol*, **18**, 577–96.

21

Genetic anemias

W. G. Wood and D. R. Higgs

Introduction

Hemoglobin, the oxygen transporting molecule that makes up over 95% of the protein content of the red blood cell, has been at the forefront of research into the genetic causes of disease since the inception of such studies in the mid-twentieth century. In this short space of time, the normal structure and function of the molecule has been described in exquisite detail and accurate structure–function relationships have been derived from analyses of structurally abnormal hemoglobins. The thalassemias demonstrated a different class of disease-causing mutations, those resulting in defects in the synthesis of the globin polypeptides. Globin genes were among the first to be cloned and sequenced, providing an extensive list of mutations while functional studies of these genes have contributed enormously to our understanding of gene regulation. Application of this knowledge through prenatal diagnosis has made a major clinical impact on a growing number of populations and progress is being made in the development of gene therapy for the severe hemoglobin disorders. All in all, hemoglobin has provided a model system for the study of human genetic diseases.

Normal human adult hemoglobin (HbA, $\alpha_2\beta_2$) is a tetrameric molecule consisting of two α-globin chains (141 amino acids) attached to two β-globin chains (146 amino acids). Each chain carries a prosthetic heme group that reversibly binds oxygen. The three dimensional structure of hemoglobin has been determined to a resolution of 1.7Å (Fermi *et al.*, 1984), demonstrating a compact elliptical shape composed largely of α-helices.

Hemoglobin provided the first example of a structurally abnormal protein as the molecular basis for a human disease, when Pauling found that hemoglobin from a patient with sickle cell anemia had a different electrophoretic mobility from HbA (Pauling *et al.*, 1949). The basis of this abnormality, a single amino acid substitution in the β chain (β6, Glu Val, HbS) was soon elucidated (Ingram, 1956). Since those early days, more than 750 abnormal hemoglobins have been described (Huisman *et al.*, 1998), the majority occurring in the heterozygous state in just one or a few families. Most are of little or no clinical consequence although variants with altered O_2 affinity, instability or formation of methemoglobin will require treatment. Three mutants that cause disease, however, have reached gene frequencies of over 10%; HbS and HbC in parts of Africa and HbE in South East Asia. In these populations, homozygotes are relatively common, which in the cases of Hbs C and E results in a fairly mild anemia, for hemoglobin S it produces a more severe disease that can be life-threatening without treatment. Moreover, these three abnormal hemoglobins occur in populations that also have a high frequency of β thalassemia (see below) and compound heterozygotes for the two conditions also result in

severe anemia, particularly in the case of HbE/β thalassemia.

The thalassemias form the other major group of genetic anemias and are due to deficiencies in the amount of either α or β globin chains synthesized. These may be either partial (α^+ or β^+) or complete (α^0 or β^0). The globin chain imbalance that results from the deficit of one of the chains is detrimental to the survival of the red cell, the damage being done not by the deficiency itself but by the resulting excess of the other chain (Schrier *et al.*, 1997). Excess α chains are unstable and precipitate within the cell, binding to and damaging the membranes. Excess β chains can initially form tetramers (β_4, HbH) but these cannot release oxygen and are also unstable. Thalassemia heterozygotes are not usually clinically affected as proteolytic mechanisms within the red cell can destroy a certain degree of excess chains. When this limit is exceeded, as in homozygotes, a severe disease results. A complete deficit of α chains usually results in death in utero while homozygotes with no β chain production tend to die in early childhood. Thalassemia occurs at high frequency in a band that stretches from the Mediterranean through the Middle East, Southern Asia to South East Asia and while over 200 different mutations have been described, in most populations only a few (3–6) alleles make up the majority of cases.

The high frequencies of HbS, HbC, HbE and the α and β thalassemias in these populations are believed to be due to strong selection by malaria, particularly by *P. falciparum*. Heterozygotes for these conditions are protected from the lethal effects of falciparum malaria to a greater degree than normal individuals. Similar selective survival in the face of malaria is believed to underlie the polymorphic frequencies reached by several other genetic abnormalities of the red cell. These include G-6-PD deficiency, South East Asian ovalocytosis (band 3 gene) and hereditary elliptocytosis (protein 4.1 deficiency).

Globin genes and their regulation

The human globin genes were among the first eukaryotic genes to be cloned and sequenced. The α globin cluster is near the telomere of the short arm of chromosome 16, while the β genes occupy an interstitial position on the short arm of chromosome 11 (Deisseroth *et al.*, 1977; 1978). Within each cluster there are a number of developmentally regulated genes arranged in order of their expression during development, together with one or more pseudogenes. Each gene consists of three exons and two introns with canonical donor and acceptor splice sites.

α globin cluster

The terminal 300 kb of the short arm of chromosome 16 containing the α globin cluster (Figure 21.1) have been fully sequenced and annotated (Flint *et al.*, 1997). The region is GC-rich and is very gene dense, with most of the genes surrounding the α cluster being expressed in a wide variety of tissues and can be considered as ''housekeeping'' genes. The three functional genes are ζ, restricted to embryonic red cells and α_2 and α_1 expressed throughout development. The duplicated α genes lie 3.7 kb apart and produce identical protein products. At the RNA level, their transcripts are distinguishable, with the α_2 gene producing approximately three times as much mRNA as the α_1 gene.

Extensive studies in transfected cell lines and in transgenic mice showed that when α globin genes alone were introduced they were incapable of being expressed, suggesting that an important regulatory element must lie some distance from the genes themselves. Subsequent searches revealed an enhancer-like sequence (HS −40) lying 40 kb upstream of the ζ globin gene (Higgs *et al.*, 1990). The site of this element is marked by an erythroid-specific, DNase 1 hypersensitive site and sequence analysis revealed it to contain a mixture of binding sites for erythroid-specific and

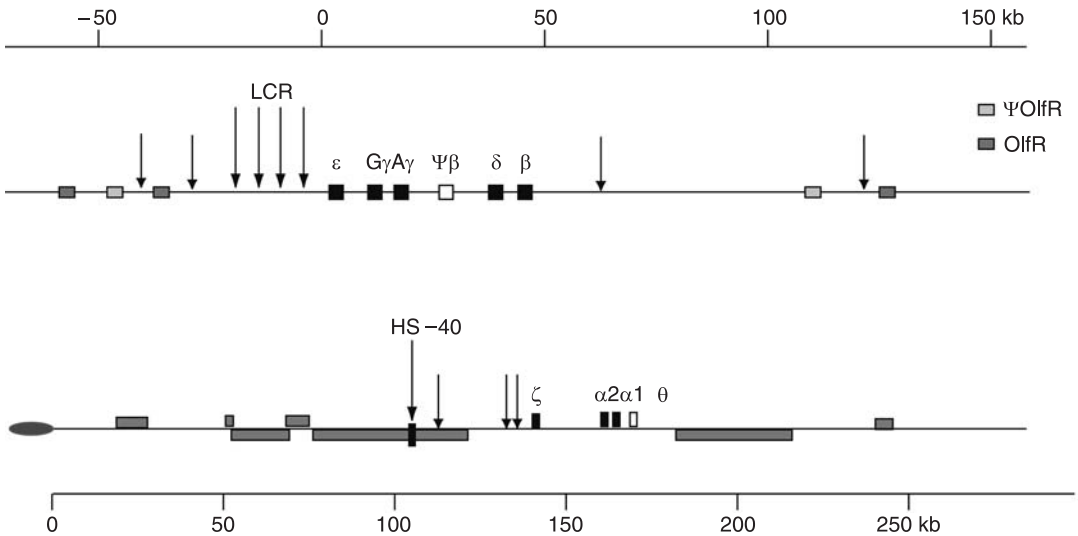

Figure 21.1 The structure of the human α and β globin gene clusters. Arrows mark the positions of DNase1 hypersensitive sites around the gene clusters and the positions of the major regulatory sequences (LCR and HS –40) are shown. Olfactory receptor genes (OlfR) and pseudogenes (ΨOlfR) flank the β globin cluster. Housekeeping genes surrounding the α genes that are transcribed in the same direction as the α genes (towards the centromere) are shown above the line while those below the line are transcribed in the opposite direction.

general transcription factors (Jarman *et al.*, 1991). Constructs containing HS −40 and the α globin genes were expressed at a high level wherever they were integrated into mouse erythroid cells (Sharpe *et al.*, 1993). Surprisingly, for an erythroid enhancer, this element lies in an intron of the upstream *C16orf35* gene which is widely expressed in most tissues (Vyas *et al.*, 1995).

Mutations of the α globin genes causing α thalassemia

Deletions of one or both α globin genes on a chromosome are the commonest cause of α thalassemia (Higgs, 2001). The original duplication that gave rise to two α globin genes has been well conserved and since these blocks lie adjacent to each other, it is not uncommon for misalignment to occur between the α1 gene of one chromosome with the α2 gene of the other, during meiosis. Crossing over would then result in either a chromosome with three α genes (which has no hematological consequences on its own) or one with a single α gene. Depending on the precise site of the cross over, the deletions differ slightly in size, with the two commonest forms losing 3.7 or 4.2 kb of sequence (depicted as $-\alpha^{3.7}$, $-\alpha^{4.2}$, respectively). These deletions are extremely common in the Mediterranean, Africa, the Middle East and Southern and South East Asia; in some populations a single α gene chromosome is more common than one with two (see below). Heterozygotes for this mild form of α thalassemia show minimal red cell changes while homozygotes have smaller red cells containing less hemoglobin than normal but are not clinically affected.

Larger deletions remove both α globin genes on the same chromosome. Most of these are rare but those found in the Mediterranean ($-^{MED}$) and in South East Asia ($-^{SEA}$) reach polymorphic frequencies. Heterozygotes for these deletions have a similar phenotype to homozygotes for the single

gene deletion, since both still have two functioning α genes. Compound heterozygotes with only one α gene (–/-α) show more severe red cell abnormalities, including the presence of HbH, and a clinical course that varies from quite mild to fairly severe anaemia. Homozygotes for the double gene deletion (–/–) are severely anemic in utero, developing hydrops fetalis and usually dying shortly before or immediately after birth.

Deletions that remove the α globin gene regulatory element but spare the structural genes themselves also cause α thalassemia. These are rare conditions in which α mRNA is reduced to <1% of normal and are important for demonstrating the essential role of the HS –40 element.

Although gene deletions are by far the most common cause of α thalassemia, point mutations may also result in this condition. Mutations that cause frameshifts, affect RNA splicing, alter the poly(A) addition signal or modify the translation initiation sequence have all been described. An additional group of mutations affect the chain termination codon of the α gene, resulting in read-through into the 3′ untranslated region of the RNA, destabilizing it. One of these termination mutants (Hb CS) is a common form of α thalassemia in South East Asia.

α thalassemia due to mutations in a *trans*-acting factor

Alpha thalassemia may rarely occur in association with developmental abnormalities including severe mental retardation, facial dysmorphism and urogenital abnormalities in an X-linked syndrome (ATRX) (Gibbons *et al.*, 1995). The *ATRX* disease gene encodes a chromatin remodeling factor of the SWI/SNF family that forms part of a large complex with a well-defined translocase activity. Although mutations in the *ATRX* gene clearly cause α thalassemia, the precise mechanism underlying this is currently unknown.

Alpha thalassemia may also occur as an acquired defect in patients with a pre-malignant condition affecting hemopoiesis called myelodysplasia (MDS) giving rise to the so called ATMDS syndrome (Higgs *et al.*, 1983). Such patients have no pre-existing, inherited form of α thalassemia but develop the condition alongside MDS. It has recently been shown that these individuals have acquired somatic mutations in the *ATRX* gene, strengthening the idea that the α genes are a direct target for the ATRX protein complex (Gibbons *et al.*, 2003).

There remain some rare families with clear evidence of α thalassemia in whom no mutations can be found in the genes or their regulatory elements and are therefore good candidates for having mutations in other unlinked loci.

β globin cluster

The β globin gene cluster on chromosome 11 is flanked by members of the olfactory receptor gene family (Figure 21.1). The five β-like genes are developmentally regulated with ε expression in the early embryo replaced largely by the paired γ genes in fetal life which in turn are succeeded by δ and β gene expression in the perinatal period. The low levels of expression of the δ gene result from base substitutions in the promoter suggesting that this gene may be evolving towards a pseudogene. As in the α cluster, high level expression of these genes requires the presence of an upstream regulatory region known as the locus control region (LCR) (Grosveld *et al.*, 1987). The LCR spans approximately 20 kb and contains four elements marked by hypersensitive sites that independently show enhancer activity in stably transfected constructs and also contain multiple transcription factor binding sites. Maximum expression levels require the whole LCR and the four HSs are believed to join together to form a holocomplex that interacts with the promoter of active genes. Two separate techniques have demonstrated that the LCR is in close proximity to the active genes in the nuclei of primary erythroid cells, but not in non-erythroid nuclei (Carter *et al.*, 2002; Tolhuis *et al.*, 2002).

Mutations of the β globin gene causing β thalassemia

Screening for heterozygous β thalassemia is initially based on the decrease in red cell indices (mean cell volume ~70fl in contrast to a normal value of ~90fl, and mean cell hemoglobin ~20pg versus ~30pg). Confirmation is obtained by measuring the proportion of the minor adult hemoglobin, HbA_2 ($\alpha_2\delta_2$) which is increased from a normal level of 2.0–3.0% to 3.5–5.5%.

In contrast to the α gene cluster, mutations causing β thalassemia are rarely gene deletions but typically point mutations affecting transcription and post-transcriptional processing of the resultant RNA. Those in which no normal β chains are produced are referred to as β^0 while those with reduced amounts of product are denoted β^+ or β^{++} thalassemia, depending on severity. The severity of the disease in compound heterozygotes or homozygotes is dependent on the particular β thalassemia alleles inherited. The milder the deficit in β chain production, the less the amount of chain imbalance and hence the milder the damage to the red cells.

Base substitutions in the CCAAC and ATAA boxes of the promoter region cause mild thalassemia, as do those in the 5′ and 3′ untranslated region. A large number of mutations cause the insertion of premature stop codons while another large group consist of insertions or deletions of one or two bases. These cause a shift in the reading frame that also tend to result in premature termination. The other large group of mutants involve RNA splicing. Mutation of the invariant donor (GT) and acceptor (AG) sites results in β^0 thalassemia, while changes to the surrounding consensus sequences produce β^+ thalassemia. Alterations to the normal sequence of either an intron or an exon may create a sequence that resembles a splice site. Use of this site may result in a marked decrease in normal RNA production. In some cases, a single base change can produce both a structurally abnormal globin and reduce the level of its output. The mutation that produces HbE, (codon 26, GAG AAG) causes an amino acid substitution (Glu Lys) but also activates a cryptic splice donor site. This reduces the amount of β^E mRNA meaning that this mutation also acts as a mild β thalassemia.

Although most β thalassemia mutations are recessive, a few have a dominant phenotype. These are largely due to frameshift mutations in exon 3. Normally, globin mRNA with a frameshift is destroyed by a mechanism of nonsense-mediated decay, but when the mutation lies close to the end of the mRNA this does not occur. Large amounts of the abnormal mRNA are available for translation and are likely to produce considerable quantities of truncated and highly unstable β globin chains. Destruction of the abnormal polypeptides produced, together with the resulting excess α chains, overwhelms the proteolytic capacity of the cell, leading to its premature destruction (Hall and Thein, 1994).

β thalassemia due to mutations in *trans*-acting factors

Although the vast majority of β thalassemia mutations are linked to the β cluster on chromosome 11, there are two rare conditions in which mutations in *trans*-acting factors cause mild forms of β thalassemia. In one family, a mutation affecting the key erythroid transcription factor GATA-1 has been shown to cause thrombocytopenia and β thalassemia (Yu *et al.*, 2002). It has been suggested that differences in the types of GATA-1 binding sites present in the α and β clusters may explain why this *trans*-acting mutation has different effects on the expression of α and β globin genes.

The second example of a trans acting mutation causing β thalassemia has been identified in patients with a rare, autosomal recessive syndrome called trichothiodystrophy characterized by dry photosensitive skin, brittle hair, short stature and variable degrees of mental retardation. Many of these individuals inherit mutations in one protein (the XPD helicase) of the multicomponent, general transcription factor TFIIH (Viprakasit *et al.*, 2001).

To date, all such patients have the hematological features of β thalassemia trait (low MCV and MCH with a raised level of HbA_2). The simplest explanation for this is that the α and β promoters have different requirements for the activity of TFIIH in initiating transcription but the details of this are currently unknown.

There are several families with β thalassemia in whom no mutations have been detected in the genes or their regulatory elements. There it seems likely that other forms of unlinked β thalassemia will be identified.

Genetic interactions

αβ thalassemia

The α and β thalassemias frequently occur at high frequencies in the same geographic areas, so it is not surprising that both conditions may be found in the same patient. It may appear paradoxical but these individuals generally run a milder course; this is because it is chain imbalance that damages the cell and a β thalassemia case with concomitant α thalassemia has less chain imbalance. Co-inheritance of heterozygous β^0 or β^+ thalassemia with α thalassemia (-α/−α or—/αα) results in more normal red cell indices, which may cause problems in population screening as their MCVs and MCHs can overlap with the normal range. In homozygous or compound heterozygous β thalassemia, the loss of one or two α genes generally results in a milder disease. Among patients clinically characterized as thalassemia intermedia rather than thalassemia major, a substantial proportion turn out to be carrying α thalassemia as well, usually -α/−α or ——/αα. This effect is more obvious in patients with milder β^+ alleles than in those with the more severe β^0 thalassemia. Conversely the inheritance of additional α globin genes exacerbates the severity of β thalassemia, such that heterozygote β thalassemics with five or six α genes usually have the phenotype of thalassemia intermedia.

Modifying genes

Homozygotes for identical β thalassemia alleles occasionally show unexpectedly mild phenotypes, suggesting that the severity of the disease can be modified by other genetic factors. One mechanism that has been demonstrated to ameliorate the disease is reactivation of the fetal γ genes. These genes are normally switched off during the perinatal period, with only a trace of HbF detectable in ~5% of adult red cells (F-cells). In homozygous β thalassemia, some HbF production persists in response to the severe anemia and due to the selective survival of F-cells, but it is insufficient to prevent disease. A number of genetic conditions, both linked and unlinked to the β globin cluster, can increase HbF production in adult life (Wood, 1993).

Large deletions that remove not only the δ and β genes but also extend up to 100 kb beyond the gene itself, result in sufficiently high levels of HbF production (15–30%) to prevent disease (hereditary persistence of fetal hemoglobin, HPFH). Similar deletions that result in lower levels (5–15%) result in a moderate disease, δβ thalassemia (Wood, 2001).

Sequence changes within the β globin cluster may also foster increased HbF production. A C → T change at position –158 of the $^{G}\gamma$ gene promoter produces a small increment in HbF in normal adults but a larger, beneficial effect in β thalassemic individuals. In other cases, genes that cause small increases in HbF (1–3%) in normal adult individuals but ameliorate the severity of β thalassemia or sickle cell disease, have been shown to be unlinked to the cluster. Such genes have been linked to chromosomes X, 6 and 8 but their function has yet to be determined (Wood, 2001).

Genotype–phenotype relationships

The interactions of the various α and β thalassemia alleles together with additional modifying genes produce a wide range of clinical consequences that can make clinical prognosis quite difficult.

As chain imbalance appears to be the major determinant of pathology, conditions with reduced imbalance will generally be milder. Thus in β thalassemia, alleles with milder deficits of β chain production will be less severe than β^0 or severe β^+ alleles. Reduction of α chain excess by co-inheriting a form of α thalassemia will also ameliorate β thalassemia as will any condition that increases γ chains to help mop up the α chain excess.

Prevalence

Thalassemia occurs at a high incidence across a broad band of the Old World, stretching from Southern Europe continuously to the Far East, as well as Northern and Central Africa (Figure 21.2). Emigrants from these regions have spread the gene worldwide. Throughout this region, the overall frequency of β thalassemia heterozygotes is around 1–5%, but this hides considerable heterogeneity. In isolated areas and island populations, frequencies of 10–20% are not uncommon, e.g. Sardinia ~12%, Cyprus, 15%. HbE (which is also a mild thalassemia) reaches heterozygote frequencies of 50–60% in parts of Laos, Cambodia and Thailand.

Although there are several hundred different β thalassemia alleles, many of them are quite rare. Each population has a limited number of common alleles, usually ~3–5, that comprise 70–90% of the cases in that region. Distributions may overlap in neighboring areas, presumably through migration but Mediterranean, African and South East Asian populations share few, if any, common alleles (Figure 21.3). This implies that each allele has

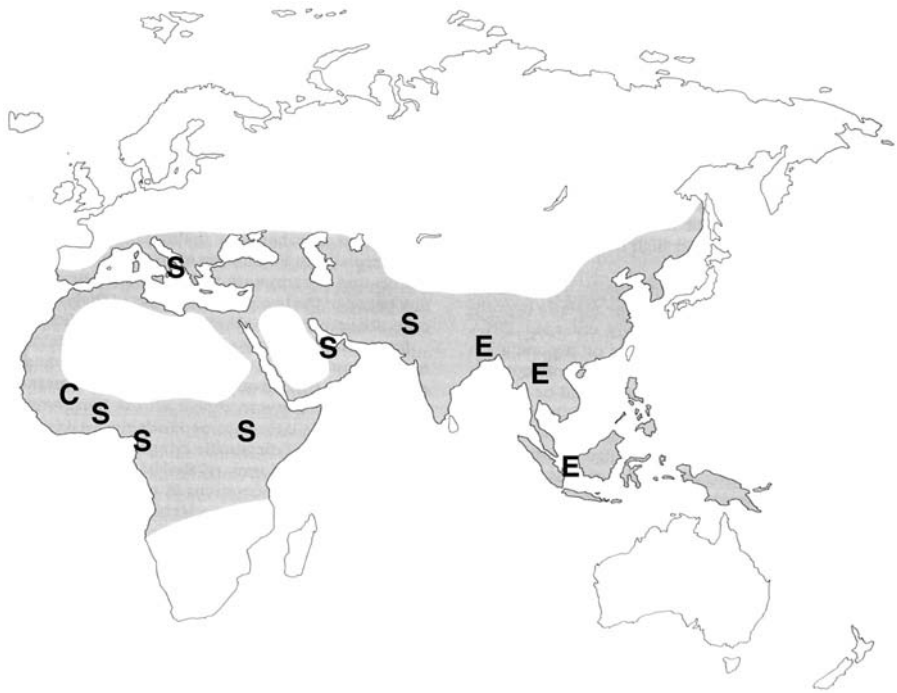

Figure 21.2 The geographical distribution of the hemoglobinopathies. The shaded region shows the distribution of the α and β thalassemias while C, S and E denote areas where HbC, HbS and HbE are prevalent (adapted from Weatherall and Clegg, 2001).

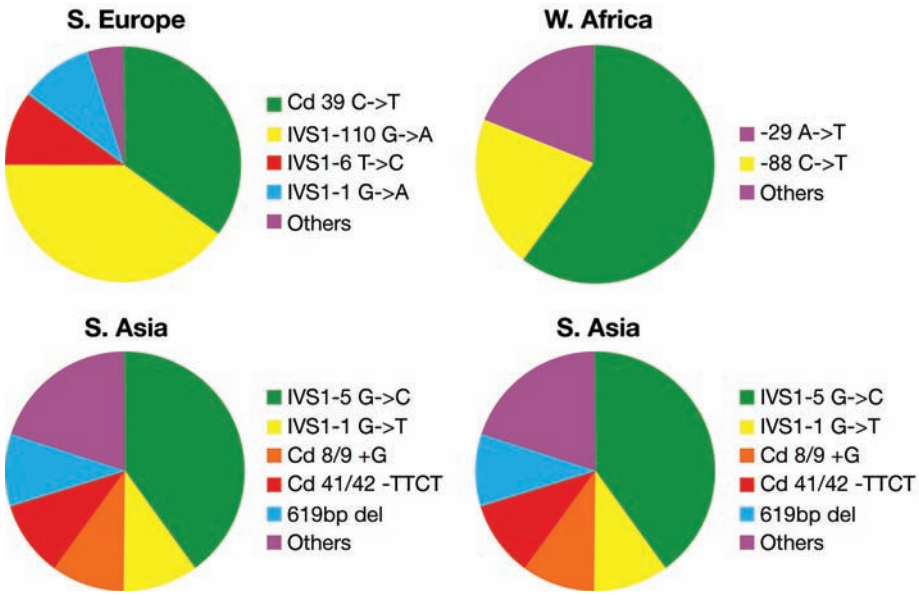

Figure 21.3 The frequencies of different β thalassemia alleles among various populations worldwide. These are average values for large areas and may show considerable variation in frequency between populations within these areas.

been expanded locally in the area it first arose. Since most alleles are largely restricted to one or two haplotypes, their expansion appears to have been relatively recent since with sufficient time, recombination would have split the haplotype association.

Of the α thalassemias, the single α gene deletion mutations ($-\alpha^{3.7}$, $-\alpha^{4.2}$) have a similar distribution to β thalassemia, whereas those with the loss of both genes (–/) have a more limited distribution, being present in the Mediterranean and South East Asia but are rare in Africa, India and the Middle East. Thus, the clinically important conditions (HbH disease and Hb Bart's hydrops fetalis) are largely limited to the former areas, where the frequency of these alleles is usually of the order of 0.1–1.0% and only in rare pockets reach 5%. In contrast, these conditions are very rare in Africa, India New Guinea and the South Pacific islands even though the -α allele often reaches frequencies of 25–50% in these areas. In some Indian tribal groups in Andra Pradesh, in southern Nepal and in parts of Papua New Guinea the -α gene frequency reaches 90%, such that the -α/-α genotype is the most common (Flint *et al.*, 1993).

Role of malaria in the selection of hemoglobinopathy alleles

The extremely high frequencies of the hemoglobinopathies in the face of the severe disease in homozygotes (and hence loss of the genes) requires explanation. Haldane (1949) first suggested that heterozygotes for these conditions may be more resistant to attacks by malaria parasites, the so-called malaria hypothesis. Evidence for this theory came initially from studies on sickle cell in Africa with the demonstration of lower parasitemia in heterozygotes and from lower deaths in childhood in AS versus AA individuals. Evidence that malaria was the selective agent in thalassemia is mainly circumstantial, but the accumulated epidemiological evidence is quite strong, particularly for α thalassemia in Melanesia (Flint *et al.*, 1986).

In vitro studies of parasite invasion and growth in hemoglobinopathic red cells have been somewhat contradictory but there is no clear evidence to date that heterozygous cells support parasite growth less well than normal cells. Nevertheless, the multiple metabolic differences in thalassemic red cells, their increased susceptibility to oxidants and altered membrane deformability suggest many ways in which they could disadvantage the parasite in vivo that may be too subtle to detect in vitro.

Public health aspects

The World Health Organization's estimate of the number of carriers of important hemoglobinopathy genes worldwide is 269 million and the estimate of newborns with a severe disease is 365 000 per year (Angastiniotis and Modell, 1998). With increasing population growth, these figures are certain to rise. Furthermore, as health standards rise in developing countries, albeit slowly, the drain on health resources due to hemoglobinopathies will become proportionately higher.

Identification of carriers of these disorders is reliable, effective and relatively inexpensive. Screening by automated blood counts and Hb electrophoresis can be applied in the antenatal clinic to detect carriers and extended to partners to identify couples at risk. This has enabled the setting up of antenatal diagnosis services in many of the more developed countries where the disorders are common. Initially, diagnosis of at risk fetuses was based on fetal blood sampling but this has now been largely replaced by direct mutation detection on DNA obtained by chorionic villus sampling. Dramatic success has been achieved in parts of the Mediterranean, Cyprus and Sardinia in particular, where births of affected children have decreased by up to 90% (Cao and Rosatelli, 1993). A major education campaign involving the general public and health professionals was essential to the success of these programmes. It remains to be seen how successfully such an approach can be transferred to the large populations of India and South East Asia.

Future treatment

At present, treatment of those patients with severe thalassemic conditions largely consists of blood transfusions to correct the anemia and treatment with iron chelating drugs, to remove the excess iron produced by transfusion. This can prolong considerably the lifespan of patients, allowing those for whom appropriate treatment is available to lead relatively normal adult lives (Olivieri and Brittenham, 1997). Iron chelation therapy is both expensive and onerous, requiring subcutaneous infusions overnight several times a week, and hence compliance rates tend to drop in adolescents. Alternative oral chelators are being developed but have yet to achieve the efficacy of desferrioxamine.

Bone marrow transplantation can cure these conditions but the proportion of patients with a matched sibling donor is low and the treatment carries an 8% mortality risk even in the best circumstances (Lucarelli *et al.*, 1995). Thus current treatment regimens, while a considerable advance, are far from satisfactory.

Prospects for gene therapy in the hemoglobinopathies have shown progress over the last few years but are still some way from application in the clinic. After rather disappointing results with retroviral vectors, due to instability of the human globin gene constructs, more promising results have been obtained using lentiviral vectors, at least in mouse models (Rivella and Sadelain, 2002). Progress in the isolation and expansion of human hemopoietic stem cells is needed, along with improvements in vector transfer efficiency to speed the development of gene therapy.

An alternative approach to those conditions that involves defective β chain production would be to reactivate the fetal γ gene, since HbF is a functioning hemoglobin in adults. A number of compounds, mostly cytotoxic drugs, are capable of increasing fetal haemoglobin in adult life

although the mechanism is not well understood. Hydroxyurea is currently in clinical use in patients with the more severe forms of sickle cell disease (Steinberg *et al.*, 2003). As HbF interferes with the formation of HbS polymers, there is less sickling and prolonged red cell survival, leading to a milder clinical disease, decreased painful crises and reduced episodes of hospitalisation. So far, hydroxyurea has not proved to be very efficacious in β thalassemia, but a few cases have responded to certain butyrate compounds with increased hemoglobin levels (Olivieri and Weatherall, 1998). A better understanding of the mechanism of the switch from HbF to HbA at birth may provide insights into how this may be reversed later in life.

What lessons can be learnt from the globin gene system for other genetic diseases?

Many principles of human genetics have been first elucidated by research focusing on the structure and expression of hemoglobin, which has always been relatively easy to study in the naturally purified population of red cells in peripheral blood. Natural mutations of hemoglobin, selected by interactions with falciparum malaria, have provided a wealth of information linking structure to function and in this enabled rapid progress in identifying the key regulatory elements located around the genes, and later more remote chromosomal elements. Very worthwhile benefits in the application of these findings to the practice of medicine have been mainly confined to population screening and prenatal diagnosis; advances in therapy for severely affected patients have been disappointing and not directly related to our understanding of underlying molecular and cellular pathophysiology. Nevertheless, in many ways, work in the globin gene system led the way forward for the application of molecular biology to medicine.

Despite these great advances it is important to consider what the ''revolution in molecular genetics'' has not explained. Perhaps the most sobering example concerns our inability to predict clinical outcome despite a seemingly complete understanding of the molecular basis for sickle cell disease, which is caused by a single base change in the β globin gene. Clearly environment plays an important role. In rural Africa, the majority of homozygotes for HbS with sickle cell disease die in infancy whereas in the more developed world the majority survive without any specific medical intervention. Many of the latter have a relatively normal life span but the severity of complications such as painful crises, stroke, acute chest complications, bone pain and necrosis are very variable. Again, the environment may have an important influence since those in better social classes have fewer complications. Even when all known interacting genetic factors (e.g. HbF expression, α thalassemia) are taken into account, we can still not predict clinical outcome in any particular individual. Today a major effort is directed towards finding new ''genetic modifiers'' of the sickle cell phenotype. But we have to ask whether we may have reached a limit using this approach and whether there may be some aspects of the natural history of human diseases which are not amenable to this type of genetic analysis either because they are due to unique combinations of genetic modifiers in each individual or because there may be some essentially random aspects to disease progression. A recent study involving a small number of identical twins with sickle cell disease who share both a common genotype and an apparently similar environment still show differences in their clinical course and endpoints. These problems highlighted by sickle cell disease are certainly seen to a lesser degree in patients with thalassemia and many diseases affecting other systems discussed elsewhere in this book.

REFERENCES

Angastiniotis, M. and Modell, B. (1998). Global epidemiology of hemoglobin disorders. *Annals of the New York Academy of Sciences*, **850**, 251–69.

Cao, A. and Rosatelli, M. C. (1993). Screening and prenatal diagnosis of the haemoglobinopathies. *Clinical Haematology*, **6**, 263–86.

Carter, D., Chakalova, L., Osborne, C. S., Dai, Y. F. and Fraser, P. (2002). Long-range chromatin regulatory interactions in vivo. *Nature Genetics*, **32**, 623–6.

Deisseroth, A., Nienhuis, A., Lawrence, J. *et al.* (1978). Chromosomal localization of human β globin gene on chromosome 11 in somatic cell hybrids. *Proceedings of the National Academy of Sciences, USA*, **75**, 1456–60.

Deisseroth, A., Nienhuis, A., Turner, P. *et al.* (1977). Localization of the human α-globin structural gene to chromosome 16 in somatic cell hybrids by molecular hybridization. *Cell*, **12**, 205–18.

Fermi, G., Perutz, M. F., Shaanan, B. and Fourme, B. (1984). The crystal structure of human deoxyhemoglobin at 1.7 Å resolution. *Journal of Molecular Biology*, **175**, 159–74.

Flint, J., Harding, R. M., Clegg, J. B. and Boyce, A. J. (1993). Why are some genetic diseases common? Distinguishing selection from other processes by molecular analysis of globin gene variants. *Human Genetics*, **91**, 91–117.

Flint, J., Hill, A. V. S., Bowden, D. K. *et al.* (1986). High frequencies of α thalassaemia are the result of natural selection by malaria. *Nature*, **321**, 744–9.

Flint, J., Thomas, K., Micklem, G. *et al.* (1997). The relationship between chromosome structure and function at a human telomeric region. *Nature Genetics*, **15**, 252–7.

Gibbons, R. J., Pellagatti, A., Garrick, D. *et al.* (2003). Identification of acquired somatic mutations in the gene encoding chromatin-remodeling factor ATRX in the alpha-thalassemia myelodysplasia syndrome (ATMDS). *Nat Genet*, **34**, 446–9.

Gibbons, R. J., Picketts, D. J., Villard, L. and Higgs, D. R. (1995). Mutations in a putative global transcriptional regulator cause X-linked mental retardation with α-thalassemia (ATR-X syndrome). *Cell*, **80**, 837–45.

Grosveld, F., van Assendelft, G. B., Breaves, D. R. and Kollias, G. (1987). Position-independent, high-level expression of the human γ-globin gene in transgenic mice. *Cell*, **51**, 975–85.

Haldane, J. B. S. (1949). The rate of mutation of human genes. Proceedings of the International Congress on VIIIth Genetics. *Hereditas*, **35**, 267–73.

Hall, G. W. and Thein, S. L. (1994). Nonsense codon mutations in the terminal exon of the β-globin gene are not associated with a reduction in β-mRNA accumulation: a mechanism for the phenotype of dominant β-thalassemia. *Blood*, **83**, 2031–7.

Higgs, D. R. (2001). α thalassaemia with mental retardation or myelodysplasia. In Weatherall, D. J. and Clegg, J. B. (eds.), *The thalassaemia syndromes*. Oxford: Blackwell, 1–846.

Higgs, D. R., Wood, W. G., Barton, C. and Weatherall, D. J. (1983). Clinical features and molecular analysis of acquired Hb H disease. *American Journal of Medicine*, **75**, 181–91.

Higgs, D. R., Wood, W. G., Jarman, A. P. *et al.* (1990). A major positive regulatory region located far upstream of the human α-globin gene locus. *Genes and Development*, **4**, 1588–601.

Huisman, T. H. J., Carver, M. F. H. and Efremov, G. D. (1998). *A syllabus of human hemoglobin variants*. Augusta, GA: Sickle Cell Foundation.

Ingram, V. M. (1956). Specific chemical difference between the globins of normal human and sickle-cell anaemia haemoglobin. *Nature*, **178**, 792–4.

Jarman, A. P., Wood, W. G., Sharpe, J. A. *et al.* (1991). Characterization of the major regulatory element upstream of the human α-globin gene cluster. *Molecular and Cellular Biology*, **11**, 4679–89.

Lucarelli, G., Giardini, C. and Baronciani, D. (1995). Bone marrow transplantation in β-thalassemia. *Seminars in Hematology*, **32**, 297–303.

Olivieri, N. F. and Brittenham, G. M. (1997). Iron-chelating therapy and the treatment of thalassemia. *Blood*, **89**, 739–61.

Olivieri, N. F. and Weatherall, D. J. (1998). The therapeutic reactivation of fetal haemoglobin. *Human Molecular Genetics*, **7**, 1655–8.

Pauling, L., Itano, H. A., Singer, S. J. and Wells, I. G. (1949). Sickle-cell anemia, a molecular disease. *Science*, **110**, 543–8.

Rivella, S. and Sadelain, M. (2002). Therapeutic globin gene delivery using lentiviral vectors. *Current Opinion of Molecular Therapy*, **4**, 505–14.

Schrier, S. L., Bunyaratvej, A., Khukapinant, A. *et al.* (1997). The unusual pathobiology of hemoglobin Constant Spring red blood cells. *Blood*, **89**, 1762–9.

Sharpe, J., Wells, D. J., Whitelaw, E. *et al.* (1993). Analysis of the human α-globin gene cluster in transgenic mice. *Proceedings of the National Academy of Sciences, USA*, **90**, 11262–6.

Steinberg, M. H., Barton, F., Castro, O. *et al.* (2003). Effect of hydroxyurea on mortality and morbidity in adult sickle cell anemia: risks and benefits up to

9 years of treatment. *Journal of the American Medical Association*, **289**, 1645–51.

Tolhuis, B., Palstra, R. J., Splinter, E., Grosveld, F. and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. *Molecular Cell*, **10**, 1453–65.

Viprakasit, V., Gibbons, R. J., Broughton, B. C. *et al.* (2001). Mutations in the general transcription factor TFIIH result in β-thalassaemia in individuals with trichothiodystrophy. *Human Molecular Genetics*, **10**, 2797–802.

Vyas, P., Vickers, M. A., Picketts, D. J. and Higgs, D. R. (1995). Conservation of position and sequence of a novel, widely expressed gene containing the major human α-globin regulatory element. *Genomics*, **29**, 679–89.

Weatherall, D. J. and Clegg, J. B. (2001). *The thalassaemia syndromes*. Oxford: Blackwell Science.

Wood, W. G. (1993). Increased Hb F in adult life. *Clinical Haematology*, **6**, 177–213.

Wood, W. G. (2001). Hereditary persistence of fetal hemoglobin and δβ thalassemia. In Forget, B. G., Higgs, D. R., Nagel, R. L. and Steinberg, M. H. eds., *Disorders of hemoglobin*, pp. 356–88. Cambridge: Cambridge University Press.

Yu, C., Niakan, K. K., Matsushita, M. *et al.* (2002). X-linked thrombocytopenia with thalassemia from a mutation in the amino finger of GATA-1 affecting DNA binding rather than FOG-1 interaction. *Blood*, **100**, 2040–5.

22

Genetics of chronic disease: obesity

I. Sadaf Farooqi and Stephen O'Rahilly

Introduction

Obesity is now recognized as one of the greatest threats to public health worldwide. The fact that its prevalence is increasing in almost every country in which reliable data is available has drawn attention to the environmental and societal factors underlying this phenomenon. However the effect of these environmental factors needs to be viewed in the light of the powerful influence of heredity on human fat mass. In this chapter we review progress in the identification of the genetic variants that influence human body fat mass and obesity.

Obesity – a major public health threat

Obesity is broadly defined as a condition in which body fat stores have increased to the extent that health may be adversely affected. Obesity is associated with substantially increased mortality from cardiovascular and cerebrovascular disease, diabetes and certain cancers (Kopelman, 2000) (Figure 22.1). It also results in morbidity from musculoskeletal, gastrointestinal, psychiatric and reproductive diseases and is associated with lowered quality of life, self-esteem and socioeconomic performance (Kopelman, 2000). The precise measurement of body fat is quite

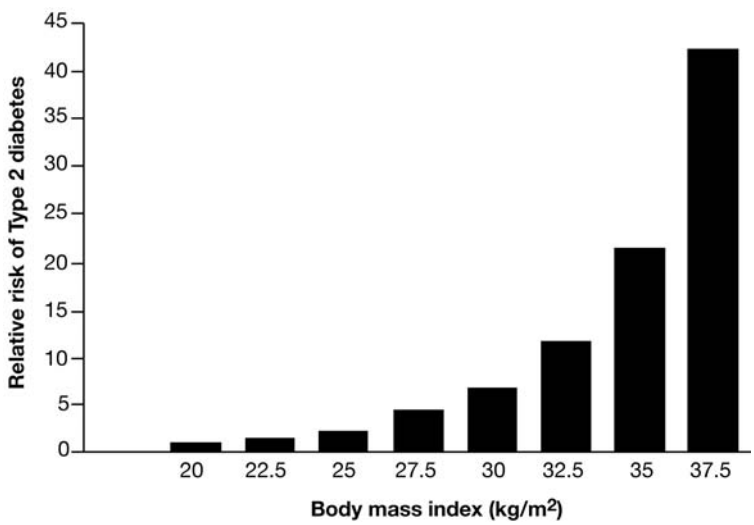

Figure 22.1 Relative risk of Type 2 diabetes with increasing BMI.

challenging and accurate methods are not applicable to large populations, therefore, surrogates such as the Body Mass Index (BMI) (weight in kilograms divided by the square of the height in metres) are most often used in population studies and in the clinic. The World Health Organization has defined obesity as a BMI exceeding 30 kg/m^2. Using this definition, currently at least 20% of the population of the USA and most Western countries is obese and at least 50% are either overweight (BMI > 25 kg/m^2) or obese. There is clear evidence, from many developing countries, of an increasing prevalence of obesity over the past two decades (Seidell, 2000). There is little doubt that the pandemic of obesity represents a major public health problem.

Is genetics important in determining susceptibility to human obesity?

The recent and relatively rapid rise in the prevalence of obesity occurring over a timescale which would make conventional evolutionary explanations implausible has, understandably, led some to question the importance of genetics in the etiology of obesity. There are some obvious candidates for the causation of secular changes in obesity prevalence including the increased availability of palatable energy dense foods and the reduced requirement for physical exertion during working and domestic life. It is therefore timely to review the evidence for inherited determinants of obesity and related intermediate traits. We will focus on data from twin and adoption studies as these are the most powerful tools to address this question.

Twin studies

The heritability of a trait is defined as the percentage of inter-individual variation in that trait which can be explained by inherited factors. Clearly heritability is not a fixed entity as the proportion of the phenotype that can be explained by the genotype will be influenced by the varying exposure to obesogenic environmental factors in different individuals. The comparison of monozygotic and dizygotic twin pairs has traditionally been one of the most powerful ways of obtaining a reliable estimate of heritability. In the case of BMI there have been over 30 published studies (Maes *et al.*, 1997) and the estimated genetic contribution to Body Mass Index ranges between 64 and 84% (Stunkard *et al.*, 1986).

Some studies have been able to bring a greater degree of sophistication to this question by studying pairs of monozygous twins reared apart as this removes the confounding factor of twins sharing very similar postnatal (if not prenatal) environments. One of the most comprehensive studies of its kind was the Swedish Adoption/Twin Study of Aging (SATSA) which included nearly 25 000 pairs of twins born in Sweden between 1886 and 1958 (Stunkard *et al.*, 1990). The intrapair correlation for monozygotic twins reared apart was the most direct estimate of heritability for BMI and was calculated as 0.70 for men and 0.66 for women. Similarly, a heritability of 0.61 was observed in a cohort of UK twins (Price and Gottesman, 1991). In a meta-analysis of results derived from Finnish, Japanese and American archival twins, Allison observed similar heritabilities (Allison *et al.*, 1996). In addition, Price and colleagues have shown that estimates of heritability did not differ significantly between twins reared apart and twins reared together, and between twins reared apart in relatively more similar (i.e. with relatives) versus less similar environments (Price and Gottesman, 1991).

Adoption studies

Adoption studies are useful in separating the common environmental effects since adoptive parents and their adoptive offspring share only environmental sources of variance, whilst the adoptees and their biological parents share only genetic sources of variance. One of the largest series, based on over 5000 subjects from the Danish adoption register which contains complete and

detailed information on the biological parents, showed a strong and statistically significant relationship between the BMI of adoptees and biological parents across the whole range of body fatness but none between adoptees and adoptive parents (Stunkard *et al.*, 1986). The Danish group have also shown a close correlation between BMI of adoptees and their biological full siblings who were reared separately by the biological parents of the adoptees, and a similar, but weaker relationship with half-siblings (Sorensen *et al.*, 1989).

Heritability of intermediate traits

In simple terms, obesity can be caused by one of three things, increased energy intake, decreased energy expenditure, or increased partitioning of nutrients into fat. These can occur alone or in combination. It would be useful if we knew the extent to which each of these processes might be under genetic control in humans. One of the serious barriers to progress in this area is the enormous difficulty in getting reliable data on human food intake in the free-living situation. In contrast, energy expenditure can be more accurately assessed using a variety of techniques, including doubly-labeled water. Bouchard and colleagues conducted some classical twin studies to examine the effects of heredity on weight gain and weight loss in response to positive and negative energy balance (Bouchard and Tremblay, 1990). In a 100-day, continually supervised inpatient overfeeding study, 12 male monozygotic twin pairs ate a 1000 kcal per day surplus (6 days per week) over the energy cost for weight maintenance (Bouchard *et al.*, 1990). The amount of weight gained during this forced overfeeding varied from 3 to 12 kgs. There was at least three times more variance in response between twin pairs than within pairs for the gains in body weight, fat mass and fat-free mass. The most powerful predictor of the amount of weight gained was the amount gained by the subject's identical twin. This strongly suggests that when food intake and exercise are controlled, inherited factors influencing either energy expenditure or nutrient partitioning have an important influence on weight gain. Similar data was obtained by the same group when inducing negative energy balance in identical twin pairs (Bouchard *et al.*, 1996). Bouchard and Tremblay have shown that about 40% of the variance in resting metabolic rate, thermic effect of food and energy cost of low to moderate intensity exercise may be explained by inherited characteristics (Bouchard and Tremblay, 1990). In addition, significant familial resemblance for level of habitual physical activity has been reported in a large cohort of healthy female twins (Samaras *et al.*, 1999). While recognizing the great difficulties inherent in the measurement of voluntary food intake, some limited twin data suggest that there are likely to be genetic influences on the overall intake of nutrients, size and frequency of meals and intake of particular foods (Wade *et al.*, 1981).

Why look for human obesity genes?

The discovery of human obesity genes will undoubtedly lead to health benefits. Firstly some obesity syndromes are very severe, occur at a young age and are associated with other developmental and clinical manifestations. It is clear that knowledge of the underlying genetic defect in these syndromes will be of considerable clinical benefit in terms of recognising other treatable aspects of the particular syndrome and providing more informed genetic counseling. As illustrated by congenital leptin deficiency (see below) the precise identification of some of these syndromes may turn out to have therapeutic relevance.

Secondly, human genetics can be an important tool to aid understanding of the molecular pathways controlling energy homeostasis in humans and thereby to clarify the relevance of such pathways as targets for pharmacological intervention.

Thirdly, and most speculatively, it is possible that by recognizing common genetic variants which predispose to obesity through different mechanisms, we can classify obese subjects into subgroups

that might have different responses to specific diets and/or exercise regimes, drugs or surgery.

Methodological issues in the hunt for human obesity genes

What measure of adiposity to use?

Obesity is defined as an excess of body fat, which can be measured directly using dual energy X-ray absorptiometry and isotopic dilution techniques. However, these are costly and their limited availability makes it difficult to perform such measurements in large numbers of subjects. In some studies fat mass has been measured indirectly using bioelectrical impedance or skin-fold thickness both of which correlate reasonably well within the normal range but less so in the very lean or the obese. The most commonly used marker of adiposity is BMI (weight in kg/height in m^2) which is a measure of heaviness that can be performed in large epidemiological studies and correlates reasonably well with body fat content. Body fat distribution provides an additional risk that is not given by simply the amount of body fat. Abdominal fat mass, for example, can vary significantly within a narrow range of total body fat or BMI. Furthermore, within a given BMI, men have on average twice the amount of abdominal fat compared to premenopausal women. There are several methods to accurately localize body fat distribution in humans (computer tomography, ultrasound and magnetic resource imaging), but these methods are often impractical and expensive and difficult to perform in large numbers of subjects. The waist to hip ratio (WHR) has been used to identify subjects with abdominal fat accumulation (WHR of >1.0 in men and >0.85 in women) as has waist circumference alone which is a convenient and simple measurement and correlates well with BMI, WHR and most importantly with risk factors for cardiovascular disease. A waist circumference of >102 cm ($\sim$40 inches) in men and >88 cm ($\sim$35 inch) in women is consistent with abdominal obesity and provides a substantial increased risk for metabolic complications. Surrogate markers of adipose tissue mass such as the adipocyte derived protein leptin (which correlates positively and closely with fat mass, $r = 0.8$) have been used in some studies (Comuzzie, 2002).

Some studies either focus on, or incorporate ''intermediate phenotypes'' in their analyses. Such traits have the theoretical advantage that they may be more proximally related to the function of the gene under study. Thus, a gene which influences energy expenditure might be easier to identify if one studied resting metabolic rate as the outcome variable. Intermediate phenotypes that are frequently used include resting metabolic rate, respiratory quotient (RQ), insulin sensitivity, and food intake and preferences measured using questionnaire-based methods.

Obesity as a disease state or adiposity as a continuous variable?

Obesity as a clinical condition is currently defined as an excess accumulation of adipose tissue resulting in a BMI greater than 30 kg/m^2. While the diagnosis of obesity by this criterion has clinical relevance with regards to intervention, management, and treatment, from an epidemiological perspective, it may hamper the study of the genetics underlying variation in body fat mass and distribution. For example, BMI exhibits a normal distribution with no clear division between the ''clinically obese'' (BMI >30) and the non-obese. Such a pattern of continuous distribution is not restricted to merely BMI but is found in all obesity-related phenotypes, including anthropometric measures (e.g. skin folds, and waist circumference), measures of body composition (e.g. percentage body fat, fat mass), and associated biochemical markers (e.g. leptin). Therefore the division of individuals into ''obese'' versus ''non-obese'' categories has a certain degree of arbitrariness that does not appear to follow any underlying biological effect (which might be due to the effect

of a gene/genes). A complex phenotype by definition is a quantifiable characteristic that is influenced by both multiple genetic and non-genetic factors as well as their interactions, leading to variation in expression across the population. Thus the use of a definition, which takes continuously distributed measures and converts them to such dichotomous states as ''obese'' vs. ''not obese,'' may actually hinder attempts to identify genes influencing complex phenotypes by greatly decreasing statistical power. Significant information may be lost in the transition from a continuous to a discontinuous scale and these issues need to be considered when selecting the appropriate method for genetic analyses of these traits (Comuzzie *et al.*, 2001).

Progress in the identification of human obesity genes

Monogenic disorders leading to human obesity

It is well established that obesity runs in families, although the vast majority of cases do not segregate with a clear Mendelian pattern of inheritance. There are about 30 Mendelian disorders with obesity as a clinical feature but often associated with mental retardation, dysmorphic features and organ-specific developmental abnormalities (i.e. pleiotropic syndromes). A number of families with these rare pleiotropic obesity syndromes have been studied by linkage analysis and the chromosomal loci for obesity syndromes are known. For a comprehensive list of syndromes in which obesity is a recognized part of the phenotype, see Online Mendelian Inheritance in Man (OMIM), www.ncbi.nlm.nih.gov/omim/.

Prader-Willi syndrome

The Prader-Willi syndrome (PWS) is the most common syndromal cause of human obesity with an estimated prevalence of about 1 in 25 000. It is an autosomal dominant disorder characterized by hypotonia, mental retardation, short stature, hypogonadotropic hypogonadism and hyperphagia (increased food intake) and obesity. Children with PWS display diminished growth, reduced lean body mass and increased fat mass – body composition abnormalities resembling those seen in growth hormone (GH) deficiency. GH treatment in these children decreases body fat, and increases linear growth, muscle mass, fat oxidation and energy expenditure (Carrel *et al.*, 1999). One suggested mediator of the obesity phenotype in PWS patients is the enteric hormone ghrelin, which is implicated in the regulation of mealtime hunger in rodents and humans and is also a potent stimulator of growth hormone secretion. Several groups have shown that children and adults with PWS have fasting plasma ghrelin levels which are 4.5-fold higher in PWS subjects than in equally obese controls and thus may be implicated in the pathogenesis of hyperphagia in these patients (Cummings *et al.*, 2002; Haqq *et al.*, 2003).

It is clear that chromosomal abnormalities are principally responsible for PWS either through deletion of the paternal ''critical'' segment 15q11.2–q12, or through loss of the entire paternal chromosome 15 with presence of 2 maternal homologues (uniparental maternal disomy). The opposite, i.e. maternal deletion or paternal uniparental disomy, causes another characteristic phenotype, the Angelman syndrome (AS). This indicates that both parental chromosomes are differentially imprinted, and that both are necessary for normal embryonic development. Most chromosomal abnormalities in PWS occur sporadically. Deletions account for 70 to 80% of cases; the majority are interstitial deletions, many of which can be visualized by prometaphase banding examination (Carrozzo *et al.*, 1997). A minority consist of unbalanced translocations, which are easily detected by routine chromosome examination. The remainder of cases are the result of maternal uniparental disomy. In most of these latter cases, cytogenetic examinations yield normal results. There are distinct differences in DNA methylation of the parental alleles at the *D15S9*

locus, providing further evidence for the association of methylation with genomic imprinting and DNA methylation can be used as a reliable postnatal diagnostic tool in PWS (Driscoll *et al.*, 1992).

Within the 4.5 Mb PWS region in 15q11–q13, where there is a lack of expression of paternally imprinted genes, several candidate genes have been studied and their expression shown to be absent in the brains of PWS patients. These include necdin, small nuclear ribonucleoprotein polypeptide N (*SNRPN*), Ring Zinc finger 127 polypeptide gene, the MAGE-like 2 gene and the Prader-Willi critical region 1 gene. However, the precise role of these genes and the mechanisms by which they lead to a pleiotropic obesity syndrome remains elusive.

Albright hereditary osteodystrophy

Albright hereditary osteodystrophy (AHO) is an autosomal dominant disorder due to germline mutations in GNAS1 that decrease expression or function of the heterotrimeric GTP binding protein, G alpha s. Maternal transmission of GNAS1 mutations leads to AHO (characterized by short stature, obesity, skeletal defects, and impaired olfaction) plus resistance to several hormones (e.g. parathyroid hormone) that activate Gs in their target tissues (pseudohypoparathyroidism type IA), while paternal transmission leads only to the AHO phenotype (pseudopseudohypoparathyroidism). Studies in both mice and humans demonstrate that GNAS1 is imprinted in a tissue-specific manner, being expressed primarily from the maternal allele in some tissues and biallelically expressed in most other tissues, thus multihormone resistance occurs only when Gs (alpha) mutations are inherited maternally (Weinstein *et al.*, 2002).

Fragile X syndrome

Fragile X syndrome is characterized by moderate to severe mental retardation, hyperactive behaviour, macro-orchidism, obesity, large ears, prominent jaw, and high-pitched jocular speech associated with mutations in the *FMR1* gene (Kaplan *et al.*, 1994). Expression is variable, with mental retardation being the most common feature. Behavioral characteristics such as hyperkinesis, autistic-like behavior, and apparent speech and language deficits may help point toward the diagnosis of the fragile X syndrome. It has been suggested that a reasonable estimate of frequency is 0.5 per 1000 males.

Bardet–Biedl syndrome

Bardet–Biedl syndrome (BBS) is a rare (prevalence < 1/100 000), autosomal recessive disease characterized by obesity, mental retardation, dysmorphic extremities (syndactyly, brachydactyly or polydactyly), retinal dystrophy or pigmentary retinopathy, hypogonadism and structural abnormalities of the kidney or functional renal impairment (Katsanis *et al.*, 2001*b*). Bardet–Biedl syndrome is a genetically heterogeneous disorder that is now known to map to at least eight loci: 11q13 (*BBS1*) (Mykytyn *et al.*, 2002); 16q21 (*BBS2*) (Nishimura *et al.*, 2001); 3p13–p12 (*BBS3*) (Sheffield *et al.*, 1994); 15q22.3–q23 (*BBS4*) (Mykytyn *et al.*, 2001); 2q31 (*BBS5*) (Young *et al.*, 1999); 20p12 (*BBS6*) (Katsanis *et al.*, 2000; Slavotinek *et al.*, 2000); 4q27 (*BBS7*) (Badano *et al.*, 2003) and 14q32.11 (*BBS8*) (Ansley *et al.*, 2003). Although BBS had been originally thought to be a recessive disorder, Katsanis and colleagues proposed that clinical manifestation of some forms of Bardet–Biedl syndrome requires recessive mutations in one of the six loci plus an additional mutation in a second locus (Katsanis *et al.*, 2001*a*). However, Mykytyn and colleagues found no evidence of involvement of the common *BBS1* mutation in triallelic inheritance (Mykytyn *et al.*, 2002).

The *BBS1* gene has recently been identified using a positional cloning approach. However, the protein does not show any significant similarity with any known protein or protein family and its function is unknown. Families with BBS mapping

to *BBS6* have been found to harbour mutations in MKKS which has sequence homology to the alpha subunit of a prokaryotic chaperonin in the thermosome *Thermoplasma acidophilum*. Mutations in this gene also cause McKusick–Kaufman syndrome (hydrometrocolpos, post-axial polydactyly and congenital heart defects). In addition, the genes underlying *BBS2* and *BBS4* have recently been identified by positional cloning. However, they have no significant homology to chaperonins and the functions of these proteins remain unknown.

Novel human monogenic obesity syndromes

In the past eight years several human disorders of energy balance that arise from genetic defects have been described. All of these are in molecules identical or similar to those known to cause obesity in genetic and experimental syndromes of obesity in rodents (Leibel *et al.*, 1997) and all have been identified using a candidate gene approach (Figure 22.2). These mutations all result in morbid obesity in childhood without the developmental pleiotropic features characteristic of the recognized syndromes of childhood obesity.

Congenital leptin deficiency

The role of leptin was first discovered in studies of severely obese *ob/ob* mice, which harbour mutations in the *ob* gene resulting in a complete lack of its protein product leptin which is derived from adipose tissue (Zhang *et al.*, 1994). Administration of recombinant leptin reduces the food intake and body weight of leptin-deficient *ob/ob* mice and corrects all their neuroendocrine and metabolic abnormalities (Halaas *et al.*, 1995).

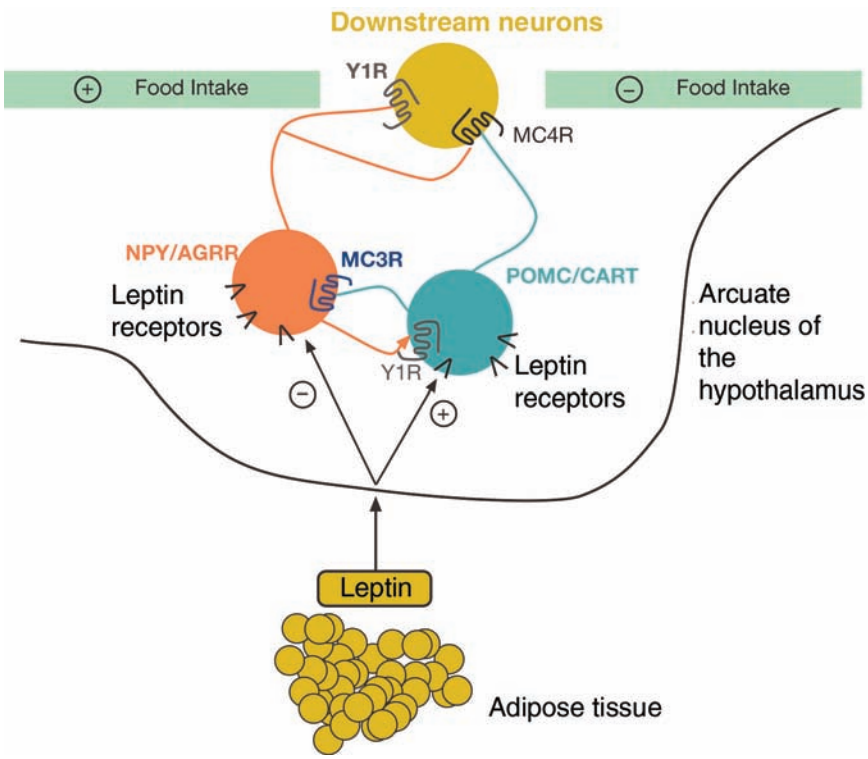

Figure 22.2 Neuronal pathways involved in the regulation of body weight.

In 1997, we reported two severely obese cousins from a highly consanguineous family of Pakistani origin who had undetectable levels of serum leptin and were found to be homozygous for a frameshift mutation in the *ob* gene (*ΔG133*), which resulted in a truncated protein that was not secreted (Montague *et al.*, 1997). We have since identified four further affected individuals from three other families who are also homozygous for the same mutation in the leptin gene. All the families are of Pakistani origin but not known to be related over five generations. A large Turkish family who carry a homozygous missense mutation have also been described (*C105T*) (Strobel *et al.*, 1998). All subjects in these families are characterised by severe early onset obesity and intense hyperphagia. Some of the Turkish subjects who are adults failed to undergo pubertal development with biochemical evidence of hypogonadotropic hypogonadism (Ozata *et al.*, 1999).

We demonstrated that children with leptin deficiency had profound abnormalities of T lymphocyte cell number and function (Farooqi *et al.*, 2002), consistent with high rates of childhood infection and a high reported rate of childhood mortality from infection in obese Turkish subjects. Most of these phenotypes closely parallel those seen in murine leptin deficiency. The contribution of reduced energy expenditure to the obesity of the *ob/ob* mouse is reasonably well established. In leptin deficient humans we found no detectable changes in resting or free-living energy expenditure, although it was not possible to examine how such systems adapted to stressors such as cold (Farooqi *et al.*, 2002). Ozata *et al.* reported abnormalities of sympathetic nerve function in leptin deficient humans consistent with defects in the efferent sympathetic limb of thermogenesis (Ozata *et al.*, 1999).

Response to leptin therapy

Recently we reported the dramatic and beneficial effects of daily subcutaneous injections of leptin reducing body weight and fat mass in three congenitally leptin-deficient children (Farooqi *et al.*, 1999; 2002). The major effect of leptin was on appetite with normalisation of hyperphagia. Leptin therapy reduced energy intake during an 18MJ ad libitum test meal by up to 84% (Farooqi *et al.*, 2002). We were unable to demonstrate a major effect of leptin on basal metabolic rate or free-living energy expenditure, but, as weight loss by other means is associated with a decrease in (BMR) basal metabolic rate, the fact that energy expenditure did not fall in our leptin-deficient subjects is notable. The administration of leptin-permitted progression of appropriately timed pubertal development in the single child of appropriate age and did not cause the early onset of puberty in the younger children (Farooqi *et al.*, 2002). Leptin also reversed the T cell dysfunction and caused a switch from a predominantly TH2 to a TH1 immune phenotype (Farooqi *et al.*, 2002).

Partial leptin deficiency in heterozygotes

We studied the heterozygous relatives of our leptin deficient subjects (Farooqi *et al.*, 2001). Serum leptin levels in the heterozygous subjects were found to be significantly lower than expected for % body fat and they had a higher prevalence of obesity than seen in a control population of similar age, sex and ethnicity (Farooqi *et al.*, 2001). Additionally, percentage body fat was higher than predicted from their height and weight in the heterozygous subjects compared to control subjects of the same ethnicity. These findings closely parallel those in heterozygous *ob-* and *db/-* mice (Chung *et al.*, 1998). These data provide further support for the possibility that leptin can produce a graded response in terms of body composition across a broad range of plasma concentrations.

Leptin receptor deficiency

The signalling form of the leptin receptor is deleted in *db/db* mice (and disrupted in the fatty Zucker and Koletsky rat models) which are consequently

unresponsive to endogenous or exogenous leptin (Tartaglia, 1997). In rodents, the phenotype is comparable to *ob/ob*, with earlier development of hyperglycemia on some backgrounds.

A mutation in the leptin receptor has been reported in several obese subjects from a consanguineous family of Kabilian origin (Clement *et al.*, 1998). Affected individuals were homozygous for a mutation that truncates the receptor before the transmembrane domain and the mutated receptor circulates bound to leptin. Although this mutation does not result in a complete null phenotype, there are a number of phenotypic similarities with the leptin-deficient subjects. Leptin receptor deficient subjects were also born of normal birth weight, exhibited rapid weight gain in the first few months of life, with severe hyperphagia and aggressive behaviour when denied food. In contrast, some neuroendocrine features were unique to leptin receptor deficiency (Clement *et al.*, 1998). The presence of mild growth retardation in early childhood with impaired basal and stimulated growth hormone secretion and decreased *IGF-1* and IGF-BP3 levels and evidence of hypothalamic hypothyroidism in these subjects, suggest that loss of the leptin receptor might result in a more diverse phenotype than loss of its ligand, leptin.

Pro-opiomelanocortin deficiency

The first-order neuronal targets of leptin action in the brain are anorectic (reducing food intake) pro-opiomelanocortin (POMC) and orexigenic (increasing food intake) neuropeptide-Y/Agouti-related protein (NPY/AgRP) neurons in the hypothalamic arcuate nucleus, where the signalling isoform of the leptin receptor is highly expressed (Schwartz *et al.*, 2000). Forty percent of POMC neurons in the arcuate nucleus express the mRNA for the long form of the leptin receptor and POMC expression is regulated positively by leptin. POMC is sequentially cleaved by prohormone convertases to yield peptides including α-MSH that have been shown to play a role in feeding behavior.

Two unrelated obese German children with homozygous or compound heterozygous mutations in POMC have been reported (Krude *et al.*, 1998). These children were hyperphagic, developing early-onset obesity as a result of impaired melanocortin signalling in the hypothalamus. They presented in neonatal life with adrenal crisis due to ACTH deficiency (POMC is a precursor of ACTH in the pituitary) and had pale skin and red hair due to the lack of MSH function at melanocortin 1 receptors in the skin (Krude *et al.*, 1998). A number of groups have identified a heterozygous missense mutation (Arg236Gly) in POMC that disrupts the dibasic amino acid processing site between β-MSH and β-endorphin, resulting in an aberrant β-MSH/β-endorphin fusion peptide which binds to *MC4R* with an affinity identical to that of α- and β-MSH but has a markedly reduced ability to activate the receptor (Challis *et al.*, 2002). Thus this cleavage site mutation in POMC may confer susceptibility to obesity through a novel molecular mechanism.

Prohormone convertase 1 deficiency

Further evidence for the role of the melanocortin system in the regulation of body weight in humans comes from the description of a 47-year-old woman with severe childhood obesity, abnormal glucose homeostasis, very low plasma insulin but elevated levels of proinsulin, hypogonadotropic hypogonadism and hypocortisolemia associated with elevated levels of POMC (O'Rahilly *et al.*, 1995). This subject was found to be a compound heterozygote for mutations in prohormone convertase 1, which cleaves prohormones at pairs of basic amino acids, leaving C-terminal basic residues which are then excised by carboxypeptidase E (CPE) (Jackson *et al.*, 1997). We have also recently identified a child with severe, early-onset obesity who was a compound heterozygote for complete loss of function mutations in PC1 (Jackson *et al.*, 2003). Although failure to cleave POMC is a likely mechanism for the obesity in these patients, PC1 cleaves a number of other neuropeptides in the

hypothalamus, such as glucagon-like-peptide 1, which may influence feeding behaviour. Intriguingly, this second patient suffered from severe small intestinal absorptive dysfunction as well as the characteristic severe early-onset obesity, impaired prohormone processing and hypocortisolemia. We hypothesized that the small intestinal dysfunction seen in this patient, and to a lesser extent in the first patient we described, may be the result of a failure of maturation of propeptides within the enteroendocrine cells and nerves that express PC1 throughout the gut. The finding of elevated levels of progastrin and proglucagon provided in vivo evidence that, indeed, prohormone processing in enteroendocrine cells was abnormal.

Melanocortin 4 receptor (*MC4R*) deficiency

There is clear evidence in rodents that α-MSH acts as a suppressor of feeding behavior probably through the melanocortin 4 receptor (*MC4R*). In fact, targeted disruption of *MC4R* in rodents leads to increased food intake, obesity, severe early hyperinsulinemia and increased linear growth; heterozygotes have an intermediate phenotype compared to homozygotes and wild-type mice (Huszar *et al.*, 1997).

Heterozygous mutations in *MC4R* have been reported to cause a dominantly inherited obesity syndrome in different ethnic groups. In a study of 500 severely obese probands, we found that approximately 5% harbored pathogenic mutations in the *MC4R* gene, making this the commonest monogenic cause of obesity thus far described in humans (Farooqi *et al.*, 2003). A small number of homozygotes for *MC4R* mutations have been described, however, the heterozygotes in these families do have an intermediate phenotype consistent with a co-dominant mode of inheritance.

Detailed phenotypic studies of patients with *MC4R* mutations reveal that this syndrome is characterised by an increase in lean body mass and bone mineral density, increased linear growth throughout childhood, hyperphagia and severe hyperinsulinemia (Farooqi *et al.*, 2003). These features are similar to those seen in *MC4R* knockout mice, suggesting the preservation of the relevant melanocortin pathways between rodents and humans. Of particular note is the finding that the severity of receptor dysfunction seen in in vitro assays can predict the amount of food ingested at a test meal by the subject harbouring that particular mutation (Farooqi *et al.*, 2003).

We have studied in detail the signaling properties of many of these mutant receptors and this information should help to advance the understanding of structure/function relationships and potentially provide in vitro support for the use of *MC4R* agonists in this group of patients (Yeo *et al.*, 2003). Importantly, we have been unable to demonstrate evidence for dominant negativity associated with these mutants, which suggests that *MC4R* mutations are more likely to result in a phenotype through haploinsufficiency.

Genetics of common obesity

Linkage studies in common obesity

The number of genome wide scans conducted in human populations has risen steadily over the last few years. Whilst many of these studies have yielded somewhat equivocal results, often due to limited sample size, there are now several studies showing evidence of linkage (LOD [likelihood of odds ratio] scores greater than 2.5) and these are discussed in this section. For a comprehensive list of all linkage studies performed see Human Obesity Gene Map (link to www.obesite.chaire.ulaval.ca/genemap.html).

To date, significant linkage results for obesity-related phenotypes have been published from a wide variety of human populations and importantly several studies have now replicated human obesity quantitative trait loci (QTLs) across different ethnic groups. To this extent there is now an emerging pattern with evidence that several major genes contribute to the variation in obesity related phenotypes in humans. Interestingly, not only are

these linkage signals supported by significant LOD scores but many of these areas contain strong positional candidate genes for obesity related phenotypes.

At present, the strongest evidence for a QTL influencing obesity-related phenotypes in humans comes from the San Antonio Family Heart Study (SAFHS) with a Log Odds Ratio (LOD) score of 7.5 for serum leptin levels on chromosome 2 (Comuzzie *et al.*, 1997). This study used a sample of 459 Mexican-Americans distributed in ten families. In addition, in this population significant linkage has also been detected on chromosome 8 with both leptin (LOD = 3.1) (Martin *et al.*, 2002) and BMI (LOD = 3.2) (Mitchell *et al.*, 1999), as well as on chromosome 17 (LOD = 3.2) for BMI (Comuzzie, 2002).

The chromosome 2 QTL localizes very near the POMC locus, which encodes pro-opiomelanocortin. Hixson and colleagues detected significant association ($P = 0.001$) between molecular variation in the POMC locus and variation in serum leptin levels among Mexican-Americans (Hixson *et al.*, 1999). The region containing the QTL on chromosome 8 contains the β-3-adrenergic receptor (*ADRB3* gene), a strong candidate previously identified on the basis of its known activity.

The Pima Indian community of Arizona, which has a particularly high prevalence of obesity, were among the first to be examined in a genome-wide scan. Currently, QTLs have been detected on chromosomes 1,3,6,11,18 and 20 with LOD scores ranging from 2 to 3.6. The strongest evidence for a QTL in the Pima Indians is on chromosome 11 (approximately 11q21–q24) with a LOD score of 3.6 for BMI (Hanson *et al.*, 1998). Additional obesity-related phenotypes, including percent body fat (LOD = 2.8) and 24-hour energy expenditure (LOD = 2.0) have also yielded linkage signals in this region in the Pima. There is evidence for a QTL located on chromosome 20 (approximately 20q11.2) for 24-hour respiratory quotient (24RQ) with a LOD score of 3.0. Norman *et al.* detected a LOD score of 2.3 on chromosome 18 (18q21) for percent body fat and a LOD score of 2.8 for 24RQ on chromosome 1 (1p31–p21) (Norman *et al.*, 1998).

Several groups have now published the results from genome screens conducted in populations of Western European (White) ancestry. The largest of these studies has been based on a collection of nuclear families collected from the mid-western region of the United States with 2209 individuals from 507 families (Kissebah *et al.*, 2000). Kissebah and colleagues detected a QTL on chromosome 3 (3q27) strongly linked to a number of traits related to obesity (including BMI, waist circumference and fasting insulin) with LOD scores ranging from 2.4 to 3.5 (Kissebah *et al.*, 2000). In addition, a second QTL influencing serum leptin levels was detected on chromosome 17 with a LOD score of 5.0. The region identified by the QTL on chromosome 3 contains several very promising positional candidate genes (e.g. adiponectin, *GLUT2*, *PI3K*), as does the region containing the chromosome 17 QTL (e.g. *GLUT4* and *PPARα*).

Hager and colleagues have published the results of a genome scan for obesity as a discrete trait (defined as a BMI > 27) in affected French sib pairs (Hager *et al.*, 1998). Significant evidence of linkage on chromosome 10 (10p) with a LOD score of 4.9 was found. The same locus showed suggestive linkage (LOD = 2.7) for BMI-adjusted leptin in an Old Order Amish population. The gene *GAD2* encoding the glutamic acid decarboxylase enzyme (GAD65) is a positional candidate gene for obesity on chromosome 10p11–12. In a recent French case-control study of 575 morbidly obese and 646 control subjects, a protective haplotype including the most frequent alleles of two single nucleotide polymorphisms (SNPs) and an at-risk SNP for morbid obesity were identified (Boutin *et al.*, 2003).

In White French sample populations, evidence of linkage to variation in serum leptin levels on chromosomes 2 and 5 has been reported. In a subsequent study of French families with Type 2 diabetes, Vionnet and colleagues reported a significant linkage signal on chromosome 3 (3q27-qter) for diabetes and glucose tolerance (Vionnet *et al.*,

2000) and evidence of linkage to diabetes on chromosomes 2 and 20.

In a genome scan of Finnish obese nuclear families, using an affected sib pair design, Ohman and colleagues identified a susceptibility locus on chromosome Xq24 (maximum likelihood score (MLS) 3.14) (Ohman *et al.*, 2000). By fine mapping, they reduced the linkage region to 4 Mb and have identified three functional candidate genes, *AGTR2*, *SLC6A14*, and *SLC25A5*. Significant evidence for association was observed for a SNP in the *SLC6A14* gene region and also with SNP haplotypes in the same gene. A significant difference in allele frequency was observed between obese and nonobese individuals in two independently ascertained Scandinavian cohorts (Suviolahti *et al.*, 2003).

Association studies in common obesity

To date, association studies have largely been restricted to candidate genes whose dysfunction might reasonably be expected to result in obesity by virtue of their having putative effects on energy intake, energy expenditure or nutrient partitioning. Genome-wide association studies are still in the future. The interpretation of association studies always requires caution because of the possibilities of problems such as population stratification, small sample size and publication bias. Some of these problems are exemplified by a common polymorphism in the β_3-adrenergic receptor, where despite over 40 association studies, involving more than 7000 subjects, the findings have been markedly inconsistent (Barsh *et al.*, 2000). While significant and consistent associations have been obtained in large case-control studies for some variants including the insulin VNTR (Le Stunff *et al.*, 2001) and *IGF-1* SNPs (Hart *et al.*, 2004), it is true to say that as of now, no single common variant is widely accepted as unequivocally associated with an alteration in human adiposity. A comprehensive and updated reference for all association studies in obesity genetics is available in the form of the obesity gene map established by Bouchard, Chagnon, Perusse and colleagues at The Pennington Biomedical Research Centre (link to www.obesite.chaire.ulaval.ca/genemap.html).

Gene/environment interaction

The increase in the prevalence of obesity in the last 30 years suggests the importance of changing environmental factors in particular the increasing availability of energy-dense, high-fat foods and a reduction in physical activity. Further evidence for the critical role of environmental factors in the development of obesity comes from migrant studies and the ''westernisation'' of diet and lifestyles in developing countries. A marked change in BMI is frequently witnessed in migrant studies, where subjects with a common genetic heritage live under new and different environmental circumstances. Pima Indians, for example, livings in the United States are on average 25 kg heavier than Pima Indians living in Mexico (Kopelman, 2000). A similar trend is seen for Africans living in the United States and Asians living in the United Kingdom. Moreover, within some ethnic groups the prevalence of obesity has increased very dramatically not only amongst migrants but also amongst the indigenous population. In fact, the prevalence of obesity is currently more than 60% in Nauruan men and women in Micronesia and amongst Polynesians in Western Samoa. These observations suggest that subjects from these ethnic groups are more susceptible to developing obesity and that environmental factors have varying effects depending upon genetic background.

The ability to identify possible interactions between genes and environmental factors is difficult because there may be a delay in an individual's exposure to an ''obesogenic'' environment and uncertainty about the precise timing of the onset of weight gain. Polymorphisms in obesity candidate genes have been studied in a few population-based cohorts on whom extensive and detailed information on diet, physical activity and markers

of intermediary metabolism have been measured. The relationship between the Pro12Ala variant in the nuclear receptor peroxisome proliferator-activated receptor-gamma (PPARgamma) and the ratio of dietary polyunsaturated fat to saturated fat (P:S ratio) has been studied and there is some evidence for a gene–nutrient interaction in a number of large studies (Lindi *et al.*, 2002; Hara *et al.*, 2000; Meirhaeghe *et al.*, 1998). Evidence for gene–exercise interactions have been found for variants in lipoprotein lipase in the HERITAGE Family Study and for the Gly16Arg variant in the beta (2)-adrenergic receptor in French and UK population-based cohorts (Meirhaeghe *et al.*, 1999; 2001). These studies, although few in number, emphasize the difficulty of examining the effect of common polymorphisms in the absence of complete data on non-genetic exposures, and may explain in part the heterogeneity of findings in previous studies.

Summary

In the last ten years, there has been considerable progress in identifying some of the genes that contribute to obesity and the regulation of body weight in humans. Although the monogenic syndromes are rare, an improved understanding of the precise nature of the inherited component of severe obesity has undoubted medical benefits for a group of patients that are most at risk from obesity-related complications. For such individuals, a genetic diagnosis helps to dispel the notion that obesity represents an individual defect in behavior with no biological basis. It also allows for informed genetic counseling and provides a starting point for providing more rational mechanism-based therapies, as has successfully been achieved for congenital leptin deficiency. Single gene disorders also serve to identify pathways in humans that are responsible for energy homeostasis and thereby clarify the relevance of such pathways as targets for pharmacological intervention.

In common polygenic obesity, the concept that environmental factors operate on an underlying pool of genes that contribute to obesity susceptibility in the general population has important implications for our approach to the prevention and treatment of obesity. Particular subsets of genetic variants may determine responses to environmental manipulations such as calorie restriction, altered macronutrient balance or exercise regimes and also response to pharmacological intervention. Thus prevention and intervention strategies can be targeted towards appropriate subgroups of the population, which is vital given the considerable number of people affected with obesity.

Whilst progress in understanding the basis for obesity as a complex trait has been slower than for other diseases, such as Type 2 diabetes, the fact that several studies are now replicating QTLs in different ethnic groups suggests that the task is probably not intractable.

REFERENCES

Allison, D. B., Kaprio, J., Korkeila, M. *et al.* (1996). The heritability of body mass index among an international sample of monozygotic twins reared apart. *Int J Obes Relat Metab Disord*, **20**, 501–6.

Ansley, S. J., Badano, J. L., Blacque, O. E. *et al.* (2003). Basal body dysfunction is a likely cause of pleiotropic Bardet–Biedl syndrome. *Nature*, **425**, 628–33.

Badano, J. L., Ansley, S. J., Leitch, C. C. *et al.* (2003). Identification of a novel Bardet–Biedl syndrome protein, BBS7, that shares structural features with BBS1 and BBS2. *Am J Hum Genet*, **72**, 650–8.

Barsh, G. S., Farooqi, I. S. and O'Rahilly, S. (2000). Genetics of body–weight regulation. *Nature*, **404**, 644–51.

Bouchard, C. and Tremblay, A. (1990). Genetic effects in human energy expenditure components. *Int J Obes*, **14** Suppl 1, 49–55; discussion 55–8.

Bouchard, C., Tremblay, A., Despres, J. P. *et al.* (1996). Overfeeding in identical twins: 5-year postoverfeeding results. *Metabolism*, **45**, 1042–50.

Bouchard, C., Tremblay, A., Despres, J. P. *et al.* (1990). The response to long-term overfeeding in identical twins. *N Engl J Med*, **322**, 1477–82.

Boutin, P., Dina, C., Vasseur, F. *et al.* (2003). GAD2 on chromosome 10p12 is a candidate gene for human obesity. *PLoS Biol*, **1**, E68.

Carrel, A. L., Myers, S. E., Whitman, B. Y. and Allen, D. B. (1999). Growth hormone improves body composition, fat utilization, physical strength and agility, and growth in Prader–Willi syndrome: a controlled study. *J Pediatr*, **134**, 215–21.

Carrozzo, R., Rossi, E., Christian, S. L. *et al.* (1997). Inter- and intrachromosomal rearrangements are both involved in the origin of 15q11–q13 deletions in Prader–Willi syndrome. *Am J Hum Genet*, **61**, 228–31.

Challis, B. G., Pritchard, L. E., Creemers, J. W. *et al.* (2002). A missense mutation disrupting a dibasic prohormone processing site in pro-opiomelanocortin (POMC) increases susceptibility to early-onset obesity through a novel molecular mechanism. *Hum Mol Genet*, **11**, 1997–2004.

Chung, W. K., Belfi, K., Chua, M. *et al.* (1998). Heterozygosity for Lep(ob) or Lep(rdb) affects body composition and leptin homeostasis in adult mice. *Am J Physiol*, **274**, R985–90.

Clement, K., Vaisse, C., Lahlou, N. *et al.* (1998). A mutation in the human leptin receptor gene causes obesity and pituitary dysfunction. *Nature*, **392**, 398–401.

Comuzzie, A. G. (2002). The emerging pattern of the genetic contribution to human obesity. *Best Pract Res Clin Endocrinol Metab*, **16**, 611–21.

Comuzzie, A. G., Hixson, J. E., Almasy, L. *et al.* (1997). A major quantitative trait locus determining serum leptin levels and fat mass is located on human chromosome 2. *Nat Genet*, **15**, 273–6.

Comuzzie, A. G., Williams, J. T., Martin, L. J. and Blangero, J. (2001). Searching for genes underlying normal variation in human adiposity. *J Mol Med*, **79**, 57–70.

Cummings, D. E., Clement, K., Purnell, J. Q. *et al.* (2002). Elevated plasma ghrelin levels in Prader Willi syndrome. *Nat Med*, **8**, 643–4.

Driscoll, D. J., Waters, M. F., Williams, C. A. *et al.* (1992). A DNA methylation imprint, determined by the sex of the parent, distinguishes the Angelman and Prader–Willi syndromes. *Genomics*, **13**, 917–24.

Farooqi, I. S., Jebb, S. A., Langmack, G. *et al.* (1999). Effects of recombinant leptin therapy in a child with congenital leptin deficiency. *N Engl J Med*, **341**, 879–84.

Farooqi, I. S., Keogh, J. M., Kamath, S. *et al.* (2001). Partial leptin deficiency and human adiposity. *Nature*, **414**, 34–5.

Farooqi, I. S., Keogh, J. M., Yeo, G. S. *et al.* (2003). Clinical spectrum of obesity and mutations in the melanocortin 4 receptor gene. *N Engl J Med*, **348**, 1085–95.

Farooqi, I. S., Matarese, G., Lord, G. M. *et al.* (2002). Beneficial effects of leptin on obesity, T cell hyporesponsiveness, and neuroendocrine/metabolic dysfunction of human congenital leptin deficiency. *J Clin Invest*, **110**, 1093–103.

Hager, J., Dina, C., Francke, S. *et al.* (1998). A genome-wide scan for human obesity genes reveals a major susceptibility locus on chromosome 10. *Nat Genet*, **20**, 304–8.

Halaas, J. L., Gajiwala, K. S., Maffei, M. *et al.* (1995). Weight-reducing effects of the plasma protein encoded by the obese gene. *Science*, **269**, 543–6.

Hanson, R. L., Ehm, M. G., Pettitt, D. J. *et al.* (1998). An autosomal genomic scan for loci linked to type II diabetes mellitus and body-mass index in Pima Indians. *Am J Hum Genet*, **63**, 1130–8.

Haqq, A. M., Farooqi, I. S., O'Rahilly, S. *et al.* (2003). Serum ghrelin levels are inversely correlated with body mass index, age, and insulin concentrations in normal children and are markedly increased in Prader–Willi syndrome. *J Clin Endocrinol Metab*, **88**, 174–8.

Hara, K., Okada, T., Tobe, K. *et al.* (2000). The Pro12Ala polymorphism in PPAR gamma2 may confer resistance to type 2 diabetes. *Biochem Biophys Res Commun*, **271**, 212–16.

Hixson, J. E., Almasy, L., Cole, S. *et al.* (1999). Normal variation in leptin levels in associated with polymorphisms in the proopiomelanocortin gene, POMC. *J Clin Endocrinol Metab*, **84**, 3187–91.

Huszar, D., Lynch, C. A., Fairchild-Huntress, V. *et al.* (1997). Targeted disruption of the melanocortin-4 receptor results in obesity in mice. *Cell*, **88**, 131–41.

Jackson, R. S., Creemers, J. W., Farooqi, I. S. *et al.* (2003). Small-intestinal dysfunction accompanies the complex endocrinopathy of human proprotein convertase 1 deficiency. *J Clin Invest*, **112**, 1550–60.

Jackson, R. S., Creemers, J. W., Ohagi, S. *et al.* (1997). Obesity and impaired prohormone processing associated with mutations in the human prohormone convertase 1 gene. *Nat Genet*, **16**, 303–6.

Kaplan, G., Kung, M., McClure, M. and Cronister, A. (1994). Direct mutation analysis of 495 patients for fragile X carrier status/proband diagnosis. *Am J Med Genet*, **51**, 501–2.

Katsanis, N., Ansley, S. J., Badano, J. L. *et al.* (2001*a*). Triallelic inheritance in Bardet–Biedl syndrome, a Mendelian recessive disorder. *Science*, **293**, 2256–9.

Katsanis, N., Beales, P. L., Woods, M. O. *et al.* (2000). Mutations in MKKS cause obesity, retinal dystrophy and renal malformations associated with Bardet–Biedl syndrome. *Nat Genet*, **26**, 67–70.

Katsanis, N., Lupski, J. R. and Beales, P. L. (2001*b*). Exploring the molecular basis of Bardet–Biedl syndrome. *Hum Mol Genet*, **10**, 2293–9.

Kissebah, A. H., Sonnenberg, G. E., Myklebust, J. *et al.* (2000). Quantitative trait loci on chromosomes 3 and 17 influence phenotypes of the metabolic syndrome. *Proc Natl Acad Sci USA*, **97**, 14478–83.

Kopelman, P. G. (2000). Obesity as a medical problem. *Nature*, **404**, 635–43.

Krude, H., Biebermann, H., Luck, W. *et al.* (1998). Severe early-onset obesity, adrenal insufficiency and red hair pigmentation caused by POMC mutations in humans. *Nat Genet*, **19**, 155–7.

Le Stunff, C., Fallin, D. and Bougneres, P. (2001). Paternal transmission of the very common class I INS VNTR alleles predisposes to childhood obesity. *Nat Genet*, **29**, 96–9.

Leibel, R. L., Chung, W. K. and Chua, S. C., Jr. (1997). The molecular genetics of rodent single gene obesities. *J Biol Chem*, **272**, 31937–40.

Lindi, V. I., Uusitupa, M. I., Lindstrom, J. *et al.* (2002). Association of the Pro12Ala polymorphism in the PPAR-gamma2 gene with 3-year incidence of type 2 diabetes and body weight change in the Finnish Diabetes Prevention Study. *Diabetes*, **51**, 2581–6.

Maes, H. H., Neale, M. C. and Eaves, L. J. (1997). Genetic and environmental factors in relative body weight and human adiposity. *Behav Genet*, **27**, 325–51.

Martin, L. J., Cole, S. A., Hixson, J. E. *et al.* (2002). Genotype by smoking interaction for leptin levels in the San Antonio Family Heart Study. *Genet Epidemiol*, **22**, 105–15.

Meirhaeghe, A., Fajas, L., Helbecque, N. *et al.* (1998). A genetic polymorphism of the peroxisome proliferator-activated receptor gamma gene influences plasma leptin levels in obese humans. *Hum Mol Genet*, **7**, 435–40.

Meirhaeghe, A., Helbecque, N., Cottel, D. and Amouyel, P. (1999). Beta2-adrenoceptor gene polymorphism, body weight, and physical activity. *Lancet*, **353**, 896.

Meirhaeghe, A., Luan, J., Selberg-Franks, P. *et al.* (2001). The effect of the Gly16Arg polymorphism of the beta(2)-adrenergic receptor gene on plasma free fatty acid levels is modulated by physical activity. *J Clin Endocrinol Metab*, **86**, 5881–7.

Mitchell, B. D., Cole, S. A., Comuzzie, A. G. *et al.* (1999). A quantitative trait locus influencing BMI maps to the region of the beta-3 adrenergic receptor. *Diabetes*, **48**, 1863–7.

Montague, C. T., Farooqi, I. S., Whitehead, J. P. *et al.* (1997). Congenital leptin deficiency is associated with severe early-onset obesity in humans. *Nature*, **387**, 903–8.

Mykytyn, K., Braun, T., Carmi, R. *et al.* (2001). Identification of the gene that, when mutated, causes the human obesity syndrome BBS4. *Nat Genet*, **28**, 188–91.

Mykytyn, K., Nishimura, D. Y., Searby, C. C. *et al.* (2002). Identification of the gene (*BBS1*) most commonly involved in Bardet–Biedl syndrome, a complex human obesity syndrome. *Nat Genet*, **31**, 435–8.

Nishimura, D. Y., Searby, C. C., Carmi, R. *et al.* (2001). Positional cloning of a novel gene on chromosome 16q causing Bardet–Biedl syndrome (BBS2). *Hum Mol Genet*, **10**, 865–74.

Norman, R. A., Tataranni, P. A., Pratley, R. *et al.* (1998). Autosomal genomic scan for loci linked to obesity and energy metabolism in Pima Indians. *Am J Hum Genet*, **62**, 659–68.

O'Rahilly, S., Gray, H., Humphreys, P. J. *et al.* (1995). Brief report: impaired processing of prohormones associated with abnormalities of glucose homeostasis and adrenal function. *N Engl J Med*, **333**, 1386–90.

Ohman, M., Oksanen, L., Kaprio, J. *et al.* (2000). Genome-wide scan of obesity in Finnish sibpairs reveals linkage to chromosome Xq24. *J Clin Endocrinol Metab*, **85**, 3183–90.

Ozata, M., Ozdemir, I. C. and Licinio, J. (1999). Human leptin deficiency caused by a missense mutation: multiple endocrine defects, decreased sympathetic tone, and immune system dysfunction indicate new targets for leptin action, greater central than peripheral resistance to the effects of leptin, and spontaneous correction of leptin-mediated defects. *J Clin Endocrinol Metab*, **84**, 3686–95.

Price, R. A. and Gottesman, II (1991). Body fat in identical twins reared apart: roles for genes and environment. *Behav Genet*, **21**, 1–7.

Samaras, K., Kelly, P. J., Chiano, M. N., Spector, T. D. and Campbell, L. V. (1999). Genetic and environmental influences on total-body and central abdominal fat: the effect of physical activity in female twins. *Ann Intern Med*, **130**, 873–82.

Schwartz, M. W., Woods, S. C., Porte, D., Jr., Seeley, R. J. and Baskin, D. G. (2000). Central nervous system control of food intake. *Nature*, **404**, 661–71.

Seidell, J. C. (2000). Obesity, insulin resistance and diabetes–a worldwide epidemic. *Br J Nutr*, **83** Suppl 1, S5–8.

Sheffield, V. C., Carmi, R., Kwitek-Black, A. *et al.* (1994). Identification of a Bardet–Biedl syndrome locus on chromosome 3 and evaluation of an efficient approach to homozygosity mapping. *Hum Mol Genet*, **3**, 1331–5.

Slavotinek, A. M., Stone, E. M., Mykytyn, K. *et al.* (2000). Mutations in MKKS cause Bardet–Biedl syndrome. *Nat Genet*, **26**, 15–16.

Sorensen, T. I., Price, R. A., Stunkard, A. J. and Schulsinger, F. (1989). Genetics of obesity in adult adoptees and their biological siblings. *Br Med J*, **298**, 87–90.

Strobel, A., Issad, T., Camoin, L., Ozata, M. and Strosberg, A. D. (1998). A leptin missense mutation associated with hypogonadism and morbid obesity. *Nat Genet*, **18**, 213–15.

Stunkard, A. J., Foch, T. T. and Hrubec, Z. (1986). A twin study of human obesity. *Jama*, **256**, 51–4.

Stunkard, A. J., Harris, J. R., Pedersen, N. L. and McClearn, G. E. (1990). The body-mass index of twins who have been reared apart. *N Engl J Med*, **322**, 1483–7.

Suviolahti, E., Oksanen, L. J., Ohman, M. *et al.* (2003). The SLC6A14 gene shows evidence of association with obesity. *J Clin Invest*, **112**, 1762–72.

Hart, L. M., Fritsche, A., Rietveld, I. *et al.* (2004). Genetic factors and insulin secretion: gene variants in the IGF genes. *Diabetes*, **53** Suppl 1, S26–30.

Tartaglia, L. A. (1997). The leptin receptor. *J Biol Chem*, **272**, 6093–6.

Vionnet, N., Hani El, H., Dupont, S. *et al.* (2000). Genomewide search for type 2 diabetes-susceptibility genes in French whites: evidence for a novel susceptibility locus for early-onset diabetes on chromosome 3q27-qter and independent replication of a type 2-diabetes locus on chromosome 1q21-q24. *Am J Hum Genet*, **67**, 1470–80.

Wade, J., Milner, J. and Krondl, M. (1981). Evidence for a physiological regulation of food selection and nutrient intake in twins. *Am J Clin Nutr*, **34**, 143–7.

Weinstein, L. S., Chen, M. and Liu, J. (2002). Gs(alpha) mutations and imprinting defects in human disease. *Ann N Y Acad Sci*, **968**, 173–97.

Yeo, G. S., Lank, E. J., Farooqi, I. S. *et al.* (2003). Mutations in the human melanocortin-4 receptor gene associated with severe familial obesity disrupts receptor function through multiple molecular mechanisms. *Hum Mol Genet*, **12**, 561–74.

Young, T. L., Penney, L., Woods, M. O. *et al.* (1999). A fifth locus for Bardet–Biedl syndrome maps to chromosome 2q31. *Am J Hum Genet*, **64**, 900–4.

Zhang, Y., Proenca, R., Maffei, M. *et al.* (1994). Positional cloning of the mouse obese gene and its human homologue. *Nature*, **372**, 425–32.

23

Type 2 diabetes mellitus

Mark I. McCarthy

Introduction

Type 2 diabetes accounts for the overwhelming majority of diabetes worldwide (Zimmet *et al.*, 2001) and represents a major and growing challenge to biomedical care. In contrast to many other complex traits, the environmental exposures which contribute to the development of this condition are well characterized: but as they are so pervasive, efforts to reduce the prevalence of this condition through environmental and behavioral manipulation have had only limited impact. Personal risk of developing type 2 diabetes results from the interaction between these pervasive exposures and our individual portfolios of susceptibility and protective genomic variants. Over the past decade, more and more of these variants have been identified and characterized. The challenge for the next decade is to understand how these variants interact with each other and with environment, and to use this information to target preventive and therapeutic interventions to maximize their effect.

Type 2 diabetes: the next global epidemic?

Definitions

In contrast to type 1 diabetes, which is known to result from autoimmune destruction of the insulin-secreting beta-cells of the pancreas, leading to lifelong dependence on exogenous insulin, the etiology of type 2 diabetes is poorly understood (Kahn, 2003). Whilst type 1 diabetes is typically diagnosed in childhood or early adulthood, type 2 diabetes classically presents in later life. These clinical distinctions lie behind previous disease classifications in which type 2 diabetes was known originally as maturity-onset diabetes, and subsequently, as non-insulin-dependent diabetes mellitus (World Health Organization Study Group, 1985).

Under the most recent classification of diabetes subtypes, published in 1997 by a joint expert committee of the American Diabetes Association and the World Health Organization, individuals whose diabetes has an established basis – be that causal genetic variation (as in some monogenic forms of diabetes (see below)) or discrete environment risk factors (e.g. certain drugs and poisons) – are excluded from the ''type 2 diabetes'' label, even if there is considerable clinical overlap (The Expert Committee on the Diagnosis and Classification of Diabetes Mellitus, 1997). There is, however, much to be learned about the molecular basis of type 2 diabetes in general from these rarer instances of known etiology. For the purposes of this chapter, therefore, the term ''type 2 diabetes'' (T2D) is used in a broad sense to include these other, related, types of diabetes.

Impact

Type 2 diabetes, and the obesity with which it is associated, represent major threats to human health in the twenty-first century. The projected increases in the prevalence of T2D are staggering.

The number of people with diabetes (most of it T2D) is expected to double in a generation, from 150 million in 2000 to 300 million by 2025 (Zimmet *et al.*, 2001). As many as one-in-ten of the people alive on the planet today will develop diabetes during their lifetime. Many hundreds of millions more will have modestly elevated blood glucose levels, which, although they do not constitute frank diabetes, are nevertheless associated with a heightened risk of cardiovascular disease (Alberti, 1996; Perry and Baron, 1999).

In the United States, the annual medical costs of caring for someone with diabetes are more than double those of a similar non-diabetic individual (Brandle *et al.*, 2003; Hogan *et al.*, 2003). This mostly reflects the costs of preventing and treating the pernicious complications of this disease, including blindness, renal failure, amputation and heart disease. Direct medical costs associated with the care of diabetes in the USA are estimated at over $90 billion, that is over 10% of total health expenditure. Yet, most of the forecast rise in diabetes prevalence will be in developing countries, where such expenditure is unsustainable. It is not hyperbole when diabetes is spoken of as the next global challenge to health after HIV/AIDS.

The gap between needs and understanding

Several landmark studies in recent years have demonstrated that the onset of T2D can be prevented in susceptible individuals by behavioral modification and the judicious use of drugs (Pan *et al.*, 1997; Tuomilehto *et al.*, 2001); and that the complications associated with T2D can be forestalled by using currently available treatments to moderate glucose levels (UK Prospective Diabetes Study Group, 1998). However, it is rare in practice for either of these approaches to realise the goal of restoring lifelong normoglycemia.

The ''classical'' model of T2D pathogenesis focuses on abnormalities in carbohydrate metabolism. This model considers that T2D results from combined defects in both insulin action (so-called insulin resistance whereby classical target tissues for insulin such as fat, muscle, and liver fail to respond appropriately to this key metabolic hormone) and insulin secretion (whereby the pancreatic beta-cell is unable to increase insulin output to overcome this resistance) (Saltiel & Kahn, 2002; Kahn, 2003). Increasingly, however, the importance of concomitant abnormalities in lipid metabolism is being appreciated (Frayn, 2001). It has been suggested for example, that the defects in insulin action and secretion are the direct consequences of obesity, raised circulating fatty acid levels, and the ectopic deposition of lipids in liver and muscle (Virkamaki *et al.*, 2001; Yki-Jarvinen, 2002).

The development of powerful new approaches to disease prevention and management is predicated on an improved molecular understanding of these fundamental abnormalities and how they lead to T2D and its complications (Moller, 2001). Given the evidence supporting a substantial genetic contribution to this disease (see below), considerable hopes have been invested in the expectation that identification of the genes influencing susceptibility, and elucidation of the ways in which these interact with the environment to generate T2D, will provide the essential clues.

Genes, environment and the pathogenesis of type 2 diabetes

The case for genes

Most researchers are nowadays comfortable with the idea that common diseases, such as T2D are complex, multifactorial traits, with individual susceptibility influenced by both genetic and environmental effects. This has not always been the case: interpretations of available data have historically been unnecessarily polarized in terms of their support for a dominant role for either ''nature'' or ''nurture''.

Evidence taken as supporting a genetic basis for T2D comes from a variety of standard

approaches. T2D clearly segregates in families (Hsueh *et al.*, 2000), and, especially in studies where allowance has been made for follow-up of apparently discordant pairs, monozygotic (identical) twins show greater pairwise concordance for glucose intolerance than do dizygotic (fraternal) pairs (Kaprio *et al.*, 1992; Poulsen *et al.*, 1999). Migration studies have repeatedly shown that individuals of South Asian origin have consistently high prevalence rates of T2D wherever they live (McKeigue *et al.*, 1989). Prevalence studies in genetically admixed populations have found that T2D rates track with the relative proportions of admixture from high- and low-risk founder populations (Brosseau *et al.*, 1979; Knowler *et al.*, 1990). Individually, however, each of these lines of evidence is amenable to alternative explanations: for example, familial aggregation may reflect shared family environment and the high prevalence of South Asian populations globally may result from shared cultural activities.

The most convincing evidence for a genetic contribution to aetiology comes from the identification of susceptibility genes. In a small proportion of pedigrees segregating T2D, clear Mendelian patterns of monogenic segregation point to penetrant etiological genes: many of these have now been identified (see below). At the same time, ongoing successes in gene discovery for more typical multifactorial T2D, including susceptibility genes mapped by both linkage and association, substantiates the view that genetic variation makes an important contribution to individual susceptibility. Neel, in his seminal ''thrifty genotype'' hypothesis (Neel, 1962), suggested that the high background prevalence of gene variants capable of increasing susceptibility to diabetes and obesity in permissive environments, reflects the fact that these same variants, favoring weight gain, fuel storage, and metabolic thrift, were positively selected during prehistory because of their survival benefit when food supply was erratic. Only in modern populations with access to unremitting calorie excess are these variants likely to be detrimental.

Heritability of quantitative traits

There is much to be gained, in most common diseases, in looking beyond the discrete disease phenotypes, and examining the quantitative traits that are thought to underlie disease development. These may reveal genetic (and environmental) influences with more clarity. In addition, the discovery of genetic variants influencing such intermediate traits may allow causal inferences to be made (Davey-Smith and Ebrahim, 2003). Relevant intermediate phenotypes for T2D include indices of insulin sensitivity and beta-cell function, and possibly, beta-cell mass (Kahn, 2003). Given the importance of obesity in T2D risk, measures of fat distribution and adipocyte function are also pertinent.

Study of these phenotypes is not particularly straightforward on two counts. Firstly, many of the traits are extremely difficult and costly to measure with any precision, especially in the large samples that may be needed for gene discovery efforts. Secondly, once individuals have established T2D, the effects of the disease itself, and/or its treatment on the values of these traits compromises their utility.

There have consequently been relatively few studies assessing the heritability of these intermediate traits (Mayer *et al.*, 1996; Elbein *et al.*, 1999*a*; Poulsen *et al.*, 1999; Watanabe *et al.*, 1999; Lehtovirta *et al.*, 2000; Hanson *et al.*, 2001). Nonetheless, the consistent picture emerges that many show substantial heritability, with indices of beta-cell function (heritabilities mostly 30–60%) generally more strongly familial than those of insulin action (20–40%).

The case for the environment

The explosive rise in T2D prevalence referred to earlier demonstrates that genetics alone cannot explain T2D susceptibility. The main drivers behind this rise are clearly the lifestyle and behavioral changes associated with Westernization and urbanization, particularly the shift towards

consumption of excessive amounts of highly processed foods, and a profound reduction in exercise and overall energy expenditure (Zimmet, 2000). Both of these promote obesity, and fuel the growing prevalence and progressively earlier onset of T2D (Ehtisham *et al.*, 2001). These effects are extremely powerful. The Pima Indians of Arizona, in whom diabetes and obesity were unknown in the early part of the twentieth century, have, within three generations, achieved the dubious claim of having the highest prevalence rates of T2D in the world with ~50% of adults diabetic (Knowler *et al.*, 1981). The same trends, only slightly less dramatically, are being repeated across the globe with increasing urbanization and industrialization. At the personal level, it is clear that regular exercise and weight loss can forestall the development of diabetes (Pan *et al.*, 1997; Tuomilehto *et al.*, 2001), once again emphasising the role of the environment in influencing disease status (though it is worth remembering that both exercise capacity and obesity are themselves partly under genetic control).

The role of early environment

The well-replicated observation that babies who are small at birth grow up to have an increased risk of T2D (and other manifestations of insulin resistance, such as heart disease and high lipids) has provided new ammunition in the gene–environment debate (Barker, 1995). The conventional explanation for this association, well supported by studies of fetal deprivation in animal models (Ozanne, 2001), emphasizes the role of adverse intra-uterine environment. Under this ''thrifty phenotype'' hypothesis, poor fetal nutrition leads not only to limited fetal growth and small size at birth, but also, through permanent shifts in metabolic performance (''metabolic programming'') to increased susceptibility to adult disease (Phillips, 1996). This effect is particularly marked when small size at birth is combined with obesity in later life. One way of viewing this is to consider programming as an appropriate metabolic adaptation to the expectation, based on fetal experience, that calorie availability will be limited; an adaptation that becomes detrimental if circumstances change.

Notwithstanding the substantial body of evidence from animal models, the evidence that this mechanism is a potent force in influencing T2D susceptibility in humans is somewhat limited. There are reports that T2D is more prevalent in individuals born following severe calorie restriction (for example during the Second World War) (Ravelli *et al.*, 1998) but these are inconsistent (Stanner *et al.*, 1997); and studies of monozygotic twins discordant for T2D have shown that the diabetic co-twin was generally the smaller one at birth (Poulsen *et al.*, 1997).

Crucially, the thrifty phenotype hypothesis also challenges the interpretation of many of the standard tools used to support the genetic basis of disease. Twin studies, for example, rely on the assumption that the environments of monozygotic twins are no more similar than those of dizygotic twins: this is questionable if intra-uterine environment is important (Phillips, 1993). Some zealous advocates of the hypothesis have suggested that early-origin effects can entirely explain T2D susceptibility, without recourse to any genetic component (Ozanne and Hales, 1998). Nevertheless, certain predictions of the thrifty phenotype hypothesis have not been sustained – for example, there is no consistent evidence for an increase in cardiovascular disease prevalence in survivors of multiple births (Christensen *et al.*, 1995) – and the early origins hypothesis has been challenged on the basis that the reported epidemiological associations reflect confounding by socioeconomic class and other factors (Kramer and Joseph, 1996).

Towards a synthesis

Other recent findings have indicated that the thrifty phenotype hypothesis cannot offer a complete explanation. An alternative (or complementary) explanation holds that observed associations between early growth and adult disease reflect

shared genetic determinants of both (Hattersley and Tooke, 1999). In several populations, it has been shown that paternal diabetes (which is known to lead to high rates of T2D in the progeny) is also associated with low offspring birth weight (Klebanoff *et al.*, 1998; Lindsay *et al.*, 2000; Hyponnen *et al.*, 2003). These observations therefore link low birth weight and subsequent T2D without recourse to maternal environmental effects. Studies in families segregating monogenic diabetes due to mutations in the glucokinase gene have found that infants inheriting these mutations are significantly smaller at birth than siblings who do not (Hattersley *et al.*, 1998). This birth weight effect is reproduced across a range of syndromes featuring severe defects in insulin secretion and action, and provides proof-of-principle that genetic variants implicated in T2D risk, also lead to lower birth weight. This is likely to reflect the crucial role played by insulin as a growth factor in early life (Hattersley and Tooke, 1999). These rare mutations cannot explain the reported population-wide associations between early growth and later disease, but suggest that common variants influencing the same mechanisms might do so. Indeed, there is some evidence, from studies of the insulin gene (Dunger *et al.*, 1998; Huxtable *et al.*, 2000), and mitochondrial genome (Casteels *et al.*, 1999), that this is the case.

Genes and environment

The hoary ''nature vs. nurture'' debate has matured in recent years. There are few who would dispute that, outside of the occasional monogenic syndrome or dramatic environmental insult, individual susceptibility to T2D is determined by the joint effects of multiple gene variants and their interaction with environment. The early-origins view reminds us that each of us is exposed to a succession of ''environments'' during our lives, that our individual genome will interact differently with each, and that the diseases we develop will reflect this cumulative and dynamic interplay. The inherent futility of trying to make precise estimates of the relative contributions of genes and environment is obvious (not just because environment is evanescent, and because the interaction of gene and environment is key) but by the fact that the fetus spends critical early months, influenced by an environment which is largely controlled by its mother, with whom it shares half its genome.

Progress towards susceptibility gene identification in type 2 diabetes

Lessons from maturity onset diabetes of the young (MODY)

As mentioned earlier, a small proportion (around 1%) of families with diabetes are clearly segregating an autosomal dominant form of early-onset T2D (Frayling *et al.*, 2001). The term ''maturity onset diabetes of the young'' (MODY) was coined to describe these key clinical features, at a time when most diabetes seen before late middle age was autoimmune (that is, type 1, or juvenile-onset diabetes as it was then known). Most people with MODY were not particularly obese, suggesting that the predominant defect in this condition was likely to be beta-cell dysfunction rather than insulin resistance. Detailed clinical studies of MODY families pointed to significant heterogeneity between families with regard to disease progression and severity (Fajans, 1990; Tattersall and Mansell, 1991).

With the advent of modern positional cloning techniques, the availability of multigenerational families and the promise of penetrant disease alleles meant that MODY became an attractive target for gene discovery efforts. Over the past decade, a total of six MODY genes have been identified, collectively explaining the molecular basis of diabetes in close to 90% of MODY pedigrees (see Table 23.1). A recent study has identified a seventh gene (*CEL*, encoding carboxyl ester lipase) that causes a MODY-like syndrome, also featuring subtle exocrine pancreatic dysfunction (Raeder *et al.*, 2006).

Table 23.1. Details of genes thus far implicated in MODY causation. Proportions of MODY families are from the UK series (Frayling *et al.*, 2001). The molecular cause of MODY is not known in about 10–15% of families

Gene ID	MODY number	Chromosomal location	Gene name	Proportion of MODY	Reference
HNF4A	*MODY1*	20q12-q13.1	Hepatocyte nuclear factor 4-alpha	2%	Yamagata *et al.*, 1996*a*
GCK	*MODY2*	7p15-p13	Glucokinase; Hexokinase 4	20%	Froguel *et al.*, 1992
HNF1A (TCF1)	*MODY3*	12q22-qter	Hepatocyte nuclear factor 1-alpha; transcription factor 1	63%	Yamagata *et al.*, 1996*b*
IPF1	*MODY4*	13q12.1	Insulin promoter factor 1; homeodomain transcription factor	<1%	Stoffers *et al.*, 1997
HNF1B (TCF2)	*MODY5*	17cen-q21.3	Hepatocyte nuclear factor 1-beta; transcription factor 2	1%	Horikawa *et al.*, 1997
NEUROD1	*MODY6*	2q32	Neurogenic differentiation1; Beta-cell E-box Transactivator 2	<1%	Malecki *et al.*, 1999

Identification of these genes has contributed significantly to our understanding of beta-cell development and function. In ~20% of MODY pedigrees, the cause is a mutation in the gene encoding the beta-cell enzyme glucokinase (Froguel *et al.*, 1992; Hattersley *et al.*, 1992; Frayling *et al.*, 2001), substantiating biochemical evidence that this enzyme was the ''glucose sensor'' of the beta-cell (Matschinsky, 1990). Haploinsufficiency at this locus leads to a resetting of this glucose sensor and a rightward shift in the glucose-stimulated insulin-response curve of the beta-cell.

In most other MODY families, diabetes results from mutations in genes encoding members of the hepatocyte nuclear factor transcription network (HNF-1α, 4α and 1β), helping to establish these as key players in the normal development and function of the beta-cell (Frayling *et al.*, 2001). The fifth and sixth genes (encoding IPF-1 and NeuroD1) are rarer but again encode transcription factors implicated in beta-cell development and function (Stoffers *et al.*, 1997; Malecki *et al.*, 1999).

This genetic heterogeneity has explained much of the clinical heterogeneity within MODY referred to earlier. The consequences of loss of one copy of the glucokinase gene are a modest rise in glucose levels, present from birth and stable with age. Since there is generally an excellent clinical response to dietary modification alone, complications are rare. In contrast, transcription-factor MODY is more severe, progressive with age, typically requires pharmacological intervention, and if inadequately treated, leads to complications. Furthermore, recent evidence from a randomised controlled trial has confirmed anecdotal evidence that individuals with *HNF-1α* mutations are acutely sensitive to the hypoglycaemic effects of sulphonylureas. This represents the first example of pharmacogenetics in diabetes care (Pearson *et al.*, 2003).

Though rare, and therefore of limited importance at the population level, the high penetrance of MODY mutations means that a molecular diagnosis has significant predictive power at the individual level (McCarthy and Hattersley, 2001). Obtaining a precise molecular diagnosis in an individual suspected to have MODY conveys not only diagnostic information (helping, for example, to confirm that a child with diabetes has MODY rather than type 1 diabetes) but also prognostic and even therapeutic insights. Reaching such a molecular diagnosis is increasingly relevant to clinical

management in early-onset T2D and MODY. However, as most MODY mutations are private, extensive screening of several genes may be needed to establish the diagnosis. With current technology, this remains prohibitively expensive for large-scale use, though this is likely to change as resequencing costs fall.

Lessons from other monogenic syndromes

Apart from MODY, T2D-like diabetes is a component of many other monogenic syndromes. These include transient neonatal diabetes mellitus, attributed to biallelic expression of an imprinted gene on chromosome 6q24 (Gardner *et al.*, 2000); and permanent neonatal diabetes mellitus, in some cases due to homozygous mutations in *GCK* (Njolstad *et al.*, 2001). Recent work has demonstrated that many cases of both transient and permanent neonatal diabetes mellitus are caused by mutations in the *KCNJ11* gene which encodes one component of the pancreatic beta-cell's K_{ATP} channel, closure of which is a key event in glucose-stimulated insulin secretion (Gloyn *et al.*, 2004; 2005). Whilst the mutated channels have lost the capacity to respond to endogenously generated adenosine triphosphate (ATP), they typically retain the ability to close in response to exogenous drugs. Many patients with neonatal diabetes have been able to dispense with lifelong insulin therapy and to achieve better, more convenient control on sulphonylureas (Zung *et al.*, 2004).

Diabetes can also result, in rare instances, from mutations in the genes encoding both insulin (leading to familial hyperproinsulinemia) and the insulin receptor (Rabson–Mendelhall syndrome). A wide range of uncommon multisystem syndromes have a prominent diabetic phenotype including Wolfram's; Alstrom's; and thiamine-responsive megaloblastic anemia (respectively attributed to mutations in *WFS1*, *ALMS1* and *SLC19A2*) (Inoue *et al.*, 1998; Diaz *et al.*, 1999; Collin *et al.*, 2002). Although rare, these diverse conditions illustrate the wide spectrum of molecular abnormalities (involving primary defects in a number of different tissues and systems) which can result in T2D. One of the most instructive of these syndromes (although not strictly a monogenic condition) is that associated with a mutation at position 3243 in the mitochondrial genome, within the gene encoding the non-nuclear form of tRNA (leucine). Depending on the tissue load of the abnormal mitochondrial sequence in any given individual, this mutation can result in either MELAS (a neurological syndrome characterized by Myalgic Encephalopathy, Lactic Acidosis and Stroke-like episodes) or Maternally-Inherited Diabetes and Deafness (MIDD) (Kadowaki *et al.*, 1994). The molecular mechanism here is thought to be one of disrupted mitochondrial metabolism, resulting in reduced oxidative phosphorylation, and ATP depletion. In the beta-cell this compromises glucose-stimulated insulin secretion.

Genome scans for linkage in multifactorial T2D

Compared with the spectacular success in identifying and characterizing the genes underlying monogenic forms of T2D, equivalent progress in the common, typical multifactorial forms of disease has been slow. This, of course, is an inevitable consequence of the etiological complexity of multifactorial disease. The rather weak correlations between any given genotype and the disease phenotype mean that gene discovery efforts are poorly powered unless very large clinical resources are deployed. Broadly speaking, as with the other complex traits described in this book, two main approaches have been adopted in gene discovery efforts (Lander and Schork, 1994). Though increasingly convergent, it is worth considering them separately, if only for historical reasons.

The genome-scan approach has sought to apply to polygenic disease, the linkage analysis tools that have proven so powerful in monogenic conditions. Over 25 genome scans for T2D linkage have been completed in a range of populations, most using collections of multiplex sibships or small pedigrees (McCarthy, 2003). Synthesizing the data emerging from these genome scans is not an easy exercise.

In T2D, there are no incontrovertible ''hits'' such as HLA in type 1 diabetes, or chromosome 16 in inflammatory bowel disease that appear repeatedly in scan after scan. Indeed, at first sight, there seems very little consistency in the locations of the major linkage signals at all. This may, in part, reflect true population differences: after all, the ''thrifty genotype'' hypothesis suggests that the metabolic variation associated with T2D may have arisen many times through convergent evolution. The most important factor is likely to be that, in the absence of a single T2D locus of overweening importance, most individual genome scans, even those including many hundreds of families, are underpowered. In such circumstances, stochastic variation plays an important part in determining which particular susceptibility loci are detected in any given study (Goring *et al.*, 2001).

Nonetheless, some chromosomal regions do seem to appear more often than expected in these scans (McCarthy, 2003). The strongest evidence exists for regions on chromosomes 20 (Ghosh *et al.*, 2000; Luo *et al.*, 2001; Permutt *et al.*, 2001; Mori *et al.*, 2002); 12q (Mahtani *et al.*, 1996; Parker *et al.*, 2001; Lindgren *et al.*, 2002); 3q (Vionnet *et al.*, 2000; Busfield *et al.*, 2002; Kissebah *et al.*, 2002; Mori *et al.*, 2002); 8p (Elbein *et al.*, 1999*b*; Wiltshire *et al.*, 2001; Busfield *et al.*, 2002); and 1q (Hanson *et al.*, 1998; Elbein *et al.*, 1999*b*; Vionnet *et al.*, 2000; Wiltshire *et al.*, 2001; Meigs *et al.*, 2002; Hsueh *et al.*, 2003). These replicated regions are obviously the most promising for further analysis, using linkage disequilibrium mapping to identify the susceptibility genes responsible.

An increasing number of promising diabetes-susceptibility variants are being uncovered through this positional cloning approach. In 1996, Hanis and colleagues (Hanis *et al.*, 1996) reported that a genome scan in Mexican-American T2D families had localized a susceptibility gene to the telomeric part of chromosome 2. Intensive linkage disequilibrium efforts subsequently refined the location of this gene, leading to evidence that the linkage effect was explained by variation in the *CAPN10* gene, encoding calpain 10, a ubiquitously expressed protease of unknown function (Horikawa *et al.*, 2000). Replicating and understanding this finding has not been easy. Firstly, the susceptibility variants implicated in the Mexican-Americans were intronic, and the strongest susceptibility effects were seen in complex models invoking both *cis*-effects and heterosis. Secondly, associations in other populations were inconsistent, and almost always weaker than in Mexican-Americans (Evans *et al.*, 2001). Thirdly, understanding of calpain biology, and its possible relationship to diabetes, has been slow to emerge. However, several recent findings are revealing strong support for the contribution of *CAPN10* variation to T2D risk. For example, in a large meta-analysis of ~7500 subjects, variation at SNP44, a variant which is in tight LD with a coding variant (T504A), was shown to be significantly associated with T2D risk (relative risk ~1.21) (Weedon *et al.*, 2003).

Recently, positional cloning efforts within the chromosome 20 region of type 2 diabetes linkage (Ghosh *et al.*, 2000) have implicated common variants within the P2 promoter of the *HNF-4 alpha* gene (one of the genes causing MODY) (Silander *et al.*, 2004; Love-Gregory *et al.*, 2004; Weedon *et al.*, 2004) in susceptibility to multifactorial diabetes. Studies in another region of linkage (chromosome 10q) have led to identification of the gene for the transcription factor *TCF7L2* as, potentially, the most important type 2 diabetes susceptibility locus yet (in terms of population attributable risk) as well as highlighting novel new aetiological pathways (involving Wnt signalling) (Grant *et al.*, 2006).

Candidate genes

The second avenue to gene discovery has been the candidate gene approach. Here, as with other complex traits, relative ignorance about the biological basis of T2D – reinforced by the evidence from monogenic syndromes that T2D can result from defects in a wide range of metabolic and cellular

processes – has often frustrated efforts to define biological candidates with high prior odds for involvement in disease susceptibility. Further, the quality of insights derived from such studies has been blighted by underpowered studies and poor attention to study design and interpretation issues (Cardon and Bell, 2001; McCarthy, 2002)

Though many hundreds of candidate genes have been studied, very few have been examined in sufficient detail as to allow any confident statements to be made concerning their potential contribution to T2D susceptibility. However, some clarity is emerging as burgeoning collaborative efforts (providing large-scale clinical resources) and improved genotyping capacity and accuracy allow adequately-powered studies to be performed.

Such analyses have identified a number of variants with strong claims to a role in T2D-susceptibility. Two of these, the P12A variant in the *PPARG* gene, and E32K in *KCNJ11* can be considered ''proven'' on the basis of cumulative evidence which now exceeds most stringent criteria for genome-wide significance.

Peroxisomal proliferative activated receptor, gamma (*PPARG*) was selected for study on the basis of its role as a key regulator of adipocyte development and function, and as the target for the thiazolidinedione group of insulin-sensitizing drugs used to treat T2D. Recently, fatty acids and eicosanoids have been identified as natural ligands for *PPARG*, reinforcing beliefs that *PPARG* plays a role in mediating the relationship between lipid metabolism and insulin action (Berger and Wagner, 2002). The initial case-control analysis of the P12A variant in this gene suggested that the minority alanine variant was associated with a substantially-decreased risk of T2D (Deeb *et al.*, 1998). Subsequent efforts at replication did not appear to confirm this finding: each study found no significant association with T2D, albeit in modest sample sizes (Altshuler *et al.*, 2000). However, studies in larger cohorts, and a meta-analysis of all available data, have provided robust evidence that the common proline allele is indeed associated with T2D (Altshuler *et al.*, 2000; Douglas *et al.*, 2001). The relative risk is modest (~1.2), but because the proline allele is common (around 85%), the population attributable risk associated with this variant may be as high as 25%.

KCNJ11 (encoding the inwardly rectifying potassium channel, Kir6.2) and its genomic neighbour *ABCC8* (encoding the sulphonylurea receptor) were strong biological candidates given that their products combine to form a critical component of the beta-cell, the membrane K_{ATP} channel (Proks *et al.*, 2002). Association studies of the E23K variant followed much the same path as those for P12A, with a clear picture only emerging once several thousands of samples had been typed (Gloyn *et al.*, 2003; Nielsen *et al.*, 2003). Again, the relative risk for T2D associated with the at-risk K allele (frequency ~40%) is modest (~1.23).

In each case, it remains unclear at this stage the sites so far implicated (P12A and E23K) are the true functional variants, or whether, because of local linkage disequilibrium, they are proxies for etiological variants elsewhere in the genes concerned (or neighbors). It is also entirely possible that susceptibility at each locus will be best described by the combined effects of multiple variants.

Other reasonable claims for susceptibility effects, though to date less convincing than those described above, exist for variants in several other genes, including those for *IRS1* (insulin receptor substrate 1) (Jellema *et al.*, 2003), *PPARGC1* (PPARG coactivator 1) (Ek *et al.*, 2001) and *INS* (insulin) (Huxtable *et al.*, 2000). There is also an increasing body of evidence that common, but less penetrant, variants in several of the genes causing MODY (*GCK*, *IPF1*, *HNF-1A*, as well as *HNF-4A*, see above) can influence susceptibility to later-onset multifactorial T2D (Stone *et al.*, 1996; MacFarlane *et al.*, 1999; Triggs-Raine *et al.*, 2002; Weedon *et al.*, 2005).

What lessons have been learnt from these candidate gene studies for T2D? The first lesson is that the relative risks associated with any individual gene are modest, though the high frequency of the at-risk alleles can mean that the population effect is substantial. Nevertheless, large sample sizes are required to demonstrate susceptibility effects, and

extremely large sample sizes will be needed to obtain precise estimates of risk, especially in the context of other genetic and environmental exposures. The second is to note that the two confirmed susceptibility variants are both in known drug targets. Not only does this inspire confidence that future susceptibility gene identification will highlight novel drug targets, it also suggests a potential for influencing treatment response. Thirdly, the genes so far implicated in T2D susceptibility are variously involved in regulating beta-cell development and function, insulin action and adipocyte function, consistent with current biological models of T2D pathogenesis. As more powerful genomics tools are applied to the understanding of T2D biology (Mootha *et al.*, 2003; Patti *et al.*, 2003), revealing genes with strong prior odds for a role in disease pathogenesis, there is every expectation that this will lead to additional susceptibility variants (McCarthy *et al.*, 2003).

Role of genetics in medical care of T2D

The value of obtaining a molecular diagnosis in individuals with early-onset forms of T2D was described earlier. The tight correlation between genotype and phenotype which characterizes MODY and other monogenic subtypes of diabetes and which has facilitated gene discovery efforts for these phenotypes, also means that genetic information gathered at these penetrant sites has the potential to deliver significant insights at the level of the individual, for example, in relation to the future risk of diabetes in family members.

Ongoing gene discovery efforts in subjects with ostensibly multifactorial T2D are likely to identify further uncommon, highly penetrant variants which define novel discrete T2D subtypes: in these cases, as with MODY, genotypes at a single locus may convey useful clinical information. However, the bulk of T2D is likely to remain irreductably multifactorial, with individual susceptibility determined by one's personal portfolio of susceptibility and protective variants at many loci, and the interaction these have with one's cumulative environmental history. In such circumstances, the clinical insights associated with scoring for any single susceptibility variant will be limited (Sotos *et al.*, 2000). For example, whilst the *PPARG* P12A variant has a population attributable risk of 25%, knowing one's genotype at this locus has only modest impact on assessment of personal risk (assuming a prior lifetime risk of 10%, risk is reduced to ~8.6% in the presence of the low-risk genotype, and rises to 10.3% if high-risk). Such subtle modulations of risk are unlikely to have much clinical value.

It is easy to show however, that much more respectable predictive power can be achieved by combining information from multiple risk factors (Pharoah *et al.*, 2002; Yang *et al.*, 2003). Obtaining even approximate estimates of the risk associated with each combination of genetic and environmental factors is not a trivial exercise. Direct measurement of risk, which is likely to preferable to use of simplistic models of interaction, will need massive, population-sized data.

For these reasons, rather than a revolution in the application of molecular methods in the management of T2D, we are much more likely to see a steady increase in usage as our understanding of the genetic contribution to susceptibility increases, and we become better able to appreciate how these factors interact with each other and environment to influence disease progression. Genetic information will be used in conjunction with other diagnostic data and information on environmental exposures to estimate and update individual profiles of disease risk, and to assess the personal costs and benefits of interventions aimed at reducing risk.

REFERENCES

Alberti, K. (1996). The clinical implications of impaired glucose tolerance. *Diabetic Medicine*, **13**, 927–37.

Altshuler, D., Hirschhorn, J. N., Klannemark, M. *et al.* (2000). The common PPARgamma Pro12Ala

polymorphism is associated with decreased risk of type 2 diabetes. *Nature Genetics*, **26**, 76–80.

Barker, D. J. P. (1995). Fetal origins of coronary heart disease. *British Medical Journal*, **311**, 171–4.

Berger, J. and Wagner, J. A. (2002). Physiological and therapeutic roles of peroxisomal proliferators-activated receptors. *Diabetes Technol Ther*, **4**, 163–74.

Brandle, M., Zhou, H., Smith, B. R. *et al.* (2003). The direct medical cost of type 2 diabetes. *Diabetes Care*, **26**, 2300–4.

Brosseau, J. D., Eelkema, R. C., Crawford, A. C. and Abe, T. A. (1979). Diabetes among the three affiliated tribes: correlation with degree of Indian inheritance. *American Journal of Public Health*, **69**, 1277–8.

Busfield, F., Duffy, D. L., Kesting, J. B. *et al.* (2002). A genomewide search for type 2 diabetes-susceptibility genes in indigenous Australians. *American Journal of Human Genetics*, **70**, 349–57.

Cardon, L. R. and Bell, J. I. (2001).Association study designs for complex diseases. *Nature Genetics Reviews*, **2**, 91–9.

Casteels, K., Ong, K., Phillips, D. *et al.* (1999). Mitochondrial 16189 variant, thinness at birth and type-2 diabetes. *Lancet*, **353**, 1499–500, (letter).

Christensen, K., Vaupel, J. W., Holm, N. V. and Yashin, A. I. (1995). Mortality among twins after age 6: fetal origins hypothesis versus twin method. *British Medical Journal*, **310**, 432–6.

Collin, G. B., Marshall, J. D., Ikeda, A. *et al.* (2002). Mutations in ALMS1 cause obesity, type 2 diabetes and neurosensory degeneration in Alstrom syndrome. *Nature Genetics*, **31**, 74–8.

Davey Smith, G. and Ebrahim, S. (2003). ''Mendelian randomization'': can genetic epidemiology contribute to understanding environmental determinants of disease? International *Journal of Epidemiology*, **32**, 1–22.

Deeb, S. S., Fajas, L., Nemoto, M. *et al.* (1998). A Pro12Ala substitution in PPARgamma2 associated with decreased receptor activity, lower body mass index and improved insulin sensitivity. *Nature Genetics*, **20**, 284–7.

Diaz, G. A., Banikazemi, M., Oishi, K., Desnick, R. J. and Gelb, B. D. (1999). Mutations in a new gene encoding a thiamine transporter cause thiamine-responsive megaloblastic anaemia syndrome. *Nature Genetics*, **22**, 309–12.

Douglas, J. A., Erdos, M. R., Watanabe, R. M. *et al.* (2001). The peroxisome proliferator-activated receptor-gamma2 Pro12Ala variant: association with type 2 diabetes and trait differences. *Diabetes*, **50**, 886–90.

Dunger, D. B., Ong, K. K. L., Huxtable, S. J. *et al.* (1998). Association of the *INS* VNTR with size at birth. *Nature Genetics*, **19**, 98–100.

Ehtisham, S., Kirk, J., McEvilly, A. *et al.* (2001). Prevalence of type 2 diabetes in children in Birmingham. *British Medical Journal*, **322**, 1428.

Ek, J., Andersen, G., Urhammer, S. A. *et al.* (2001). Mutation analysis of peroxisome proliferator-activated receptor-γ coactivator-1 (PGC-1) and relationships of identified amino acid polymorphisms to type II diabetes mellitus. *Diabetologia*, **44**, 2220–6.

Elbein, S. C., Hasstedt, S. J., Wegner, K. and Kahn, S. E. (1999*a*). Heritability of pancreatic β-cell function among nondiabetic members of Caucasian familial type 2 diabetic kindreds. *Journal of Clinical Endocrinology and Metabolism*, **84**, 1398–403.

Elbein, S. C., Hoffman, M. D., Teng, K., Leppert, M. F. and Hasstedt, S. J. (1999*b*). A genome-wide search for type 2 diabetes susceptibility genes in Utah Caucasians. *Diabetes*, **48**, 1175–82.

Evans, J. C., Frayling, T. M., Cassell, P. G. *et al.* (2001). Studies of association between the gene for calpain-10 and type 2 diabetes mellitus in the United Kingdom. *American Journal of Human Genetics*, **69**, 544–52.

Fajans, S. (1990). Scope and heterogeneous nature of MODY. *Diabetes Care*, **13**, 49–64.

Frayling, T. M., Evans, J. C., Bulman, M. P. *et al.* (2001). β-cell genes and diabetes. Molecular and clinical characterization of mutations in transcription factors. *Diabetes*, **50**, Suppl 1, S94–S100.

Frayn, K. N. (2001). Adipose tissue and the insulin resistance syndrome. *Proceedings of the Nutrition Society*, **60**, 375–80.

Froguel, Ph, Vaxillaire, M., Sun, F. *et al.* (1992). Close linkage of glucokinase locus on chromosome 7p to early-onset non-insulin-dependent diabetes mellitus. *Nature*, **356**, 162–5.

Gardner, R. J., Mackay, D. J. G., Mungall, A. J. *et al.* (2000). An imprinted locus associated with transient neonatal diabetes mellitus. *Human Molecular Genetics*, **9**, 589–96.

Ghosh, S., Watanabe, R. M., Valle, T. T. *et al.* (2000). The Finland–United States Investigation of non-insulin-dependent diabetes mellitus genetics (FUSION) study. I. An autosomal genome scan for genes that predispose to type 2 diabetes. *American Journal of Human Genetics*, **67**, 1174–85.

Gloyn, A. L., Weedon, M. N., Owen, K. R. *et al.* (2003). Large scale association studies of variants in genes encoding

the pancreatic beta-cell K-ATP channel subunits Kir6.2 (KCNJ11) and SUR1 (ABCC8) confirm that the KCNJ11 E23K variant is associated with increased risk of Type 2 Diabetes. *Diabetes*, **52**, 568–72.

Gloyn, A. L., Pearson, E. R., Antcliff, J. F. *et al.* (2004). Activating mutations in the gene encoding the ATP-sensitive potassium-channel subunit Kir6.2 and permanent neonatal diabetes. *New England Journal of Medicine*, **350**, 1838–49.

Gloyn, A. L., Reimann, F., Girard, C. *et al.* (2005). Relapsing diabetes can result from moderately activating mutations in KCNJ11. *Human Molecular Genetics*, **14**, 925–34.

Goring, H. H. H., Terwilliger, J. D. and Blangero, J. (2001). Large upward bias in estimation of locus-specific effects from genome-wide scans. *American Journal of Human Genetics*, **69**, 1357–69.

Grant, S. F., Thorleifsson, G., Reynisdottir, I. *et al.* (2006). Variant of transcription factor 7-like 2 (*TCF7L2*) gene confers risk of type 2 diabetes. *Nature Genetics*, **38**, 320–3.

Hanis, C. L., Boerwinkle, E., Chakraborty, R. *et al.* (1996). A genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome 2. *Nature Genetics*, **13**, 161–71.

Hanson, R. L., Ehm, M. G., Pettitt, D. J. *et al.* (1998). An autosomal genomic scan for loci linked to type II diabetes mellitus and body-mass index in Pima Indians. *American Journal of Human Genetics*, **63**, 1124–32.

Hanson, R. L., Imperatore, G., Narayan, K. M. *et al.* (2001). Family and genetic studies of indices of insulin sensitivity and insulin secretion in Pima Indians. *Diabetes Metabolism Research Reviews*, **17**, 296–303.

Hattersley, A. T., Turner, R. C., Permutt, M. A. *et al.* (1992). Linkage of type 2 diabetes to the glucokinase gene. *Lancet*, **339**, 1307–10.

Hattersley, A. T., Beards, F., Ballantyne, E. *et al.* (1998). Mutations in the glucokinase gene of the foetus result in reduced birth weight. *Nature Genetics*, **19**, 268–70.

Hattersley, A. T. and Tooke, J. E. (1999). The fetal insulin hypothesis an alternative explanation of the association of low birthweight with diabetes and vascular disease. *Lancet*, **353**, 1789–92.

Hogan, P., Dall, T., Nikolov, P. and American Diabetes Association (2003). Economic costs of diabetes in the US in 2002. *Diabetes Care*, **26**, 917–32.

Horikawa, Y., Iwasaki, N., Hara, M. *et al.* (1997). Mutation in hepatocyte nuclear factor-1β gene (TCF2) associated with MODY. *Nature Genetics*, **17**, 384–5.

Horikawa, Y., Oda, N., Cox, N. J. *et al.* (2000). Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. *Nature Genetics*, **26**, 163–75.

Hsueh, W. C., Mitchell, B. D., Aburomia, R. *et al.* (2000). Diabetes in the Old Order Amish: characterization and heritability analysis of the Amish Family Diabetes Study. *Diabetes Care*, **23**, 595–601.

Hsueh, W. C., St Jean, P. L., Mitchell, B. D. *et al.* (2003). Genome-wide and fine-mapping linkage studies of type 2 diabetes and glucose traits in the Old Order Amish: evidence for a new diabetes locus on chromosome 14q11 and confirmation of a locus on chromosome 1q21-q24. *Diabetes*, **52**, 550–7.

Huxtable, S. J., Saker, P. J., Haddad, L. *et al.* (2000). Analysis of parent–offspring trios provides evidence for linkage and association between the insulin gene and type 2 diabetes mediated exclusively through paternally transmitted class III variable number tandem repeat alleles. *Diabetes*, **49**, 126–30.

Hypönnen, E., Davey Smith, G. and Power, C. (2003). Parental diabetes and birth weight of offspring: intergenerational cohort study. *British Medical Journal*, **326**, 19–20.

Inoue, H., Tanizawa, Y., Wasson, J. *et al.* (1998). A gene encoding a transmembrane protein is mutated in patients with diabetes mellitus and optic atrophy (Wolfram syndrome). *Nature Genetics*, **20**, 143–8.

Jellema, A., Zeegers, M. P., Feskens, E. J., Dagnelie, P. C. and Mensink, R. P. (2003). Gly972Arg variant in the insulin receptor substrate-1 gene and association with type 2 diabetes: a metaanalysis of 27 studies. *Diabetologia*, **46**, 990–5.

Kadowaki, T., Kadowaki, H., Mori, Y. *et al.* (1994). A subtype of diabetes mellitus associated with a mutation of mitochondrial DNA. *New England Journal of Medicine*, **330**, 962–8.

Kahn, S. E. (2003). The relative contributions of insulin resistance and beta-cell dysfunction to the pathophysiology of type 2 diabetes. *Diabetologia*, **46**, 3–19.

Kaprio, J., Tuomilehto, J., Koskenvuo, M. *et al.* (1992). Concordance for type 1 (insulin-dependent) and type 2 (non-insulin-dependent) diabetes mellitus in a population-based cohort of twins in Finland. *Diabetologia*, **35**, 1060–7.

Kissebah, A. H., Sonnenberg, G. E., Myklebust, J. *et al.* (2002). Quantitative trait loci on chromosomes 3 and 17

influence phenotypes of the metabolic syndrome. *Proceedings of the National Academy of Sciences USA*, **97**, 14478–83.

Klebanoff, M. A., Mednick, B. R., Schulsinger, C., Secher, N. J. and Shiono, P. H. (1998). Father's effect on infant birth weight. *American Journal of Obstetrics and Gynaecology*, **178**, 1022–6.

Knowler, W. C., Pettitt, D. J., Saad, M. F. and Bennett, P. H. (1990). Diabetes mellitus in the Pima Indians: incidence, risk factors and pathogenesis. *Diabetes/Metabolism Reviews*, **6**, 1–27.

Knowler, W. C., Pettitt, D. J., Savage, P. J. and Bennett, P. H. (1981). Diabetes incidence in Pima Indians: contributions of obesity and parental diabetes. *American Journal of Epidemiology*, **113**, 144–56.

Kramer, M. S. and Joseph, K. S. (1996). Enigma of fetal/infant-origins hypothesis. *Lancet*, **348**, 1254–5.

Lander, E. S. and Schork, N. J. (1994). Genetic dissection of complex traits. *Science*, **265**, 2037–48.

Lehtovirta, M., Kaprio, J., Forsblom, C. *et al.* (2000). Insulin sensitivity and insulin secretion in monozygotic and dizygotic twins. *Diabetologia*, **43**, 285–93.

Lindgren, C. M., Mahtani, M. M., Widen, E. *et al.* (2002). Genomewide search for type 2 diabetes mellitus susceptibility loci in Finnish families: the Botnia Study. *American Journal of Human Genetics*, **70**, 509–16.

Lindsay, R. S., Dabelea, D., Roumain, J. *et al.* (2000). Type 2 diabetes and low birth weight: the role of paternal inheritance in the association of low birth weight and diabetes. *Diabetes*, **49**, 445–9.

Love-Gregory, L., Wasson, J., Ma, J. *et al.* (2004). A common polymorphism in the upstream promoter region of the hepatocyte nuclear factor-4 gene on chromosome 20q is associated with Type 2 Diabetes and appears to contribute to the evidence for linkage in an Ashkenazi Jewish population. *Diabetes*, **53**, 1134–40.

Luo, T. H., Zhao, Y., Li, G. *et al.* (2001). A genome-wide search for type II diabetes susceptibility genes in Chinese Hans. *Diabetologia*, **44**, 501–6.

MacFarlane, W. M., Frayling, T. M., Ellard, S. *et al.* (1999). Missense mutations in the insulin promoter factor-1 gene predispose to type 2 diabetes. *Journal of Clinical Investigation*, **104**, R33–R39.

Mahtani, M. M., Widén, E., Lehto, M. *et al.* (1996). Mapping of a gene for NIDDM associated with an insulin secretion defect by a genome scan in Finnish families. *Nature Genetics*, **14**, 90–5.

Malecki, M. T., Jhala, U. S., Antonellis, A. *et al.* (1999). Mutations in NEUROD1 are associated with the development of type 2 diabetes mellitus. *Nature Genetics*, **23**, 323–8.

Matschinsky, F. M. (1990). Glucokinase as glucose sensor and metabolic signal generator in pancreatic beta-cells and hepatocytes. *Diabetes*, **39**, 647–52.

Mayer, E. J., Newman, B., Austin, M. A. *et al.* (1996). Genetic and environmental influences on insulin levels and the insulin resistance syndrome: an analysis of women twins. *American Journal of Epidemiology*, **143**, 323–32.

McCarthy, M. I. and Hattersley, A. T. (2001). Molecular diagnostics in monogenic and multifactorial forms of type 2 diabetes. *Expert Review of Molecular Diagnostics*, **1**, 403–12.

McCarthy, M. I. (2002). Susceptibility gene discovery in common metabolic and endocrine traits. *Journal of Molecular Endocrinology*, **28**, 1–17.

McCarthy, M. I. (2003). Growing evidence for diabetes susceptibility genes from genome scan data. *Diabetes Current Reports*, **3**, 159–67.

McCarthy, M. I., Smedley, D. and Hide, W. (2003). New methods for finding disease susceptibility genes: Impact and potential. *Genome Biology*, **4**, 119.1–119.8.

McKeigue, P. M., Miller, G. J. and Marmot, M. G. (1989). Coronary heart disease in south Asians overseas: a review. *Journal of Clinical Epidemiology*, **42**, 597–609.

Meigs, J. B., Panhuysen, C. I., Myers, R. H., Wilson, P. W. and Cupples, L. A. (2002). A genome-wide scan for loci linked to plasma levels of glucose and HbA(1c) in a community-based sample of Caucasian pedigrees: the Framingham Offspring Study. *Diabetes*, **51**, 833–40.

Moller, D. E. (2001). New drug targets for type 2 diabetes and the metabolic syndrome. *Nature*, **414**, 821–7.

Mootha, V. K., Lindgren, C. M., Eriksson, K.-F. *et al.* (2003). PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature Genetics*, **34**, 267–73.

Mori, Y., Otabe, S., Dina, C. *et al.* (2002). Genome-wide search for type 2 diabetes in Japanese affected sib-pairs confirms susceptibility genes on 3q, 15q, and 20q and identifies two new candidate loci on 7p and 11p. *Diabetes*, **51**, 1247–55.

Neel, J. V. (1962). Diabetes mellitus: a ''thrifty'' genotype rendered detrimental by ''progress''? *American Journal of Human Genetics*, **14**, 353–62.

Nielsen, E. D., Hansen, L., Carstensen, B. *et al.* (2003). The E23K variant of Kir6.2 associates with impaired post-OGTT serum insulin response and increased risk of type 2 diabetes. *Diabetes*, **52**, 573–7.

Njolstad, P. R., Sovik, O., Cuesta-Munoz, A. *et al.* (2001). Neonatal diabetes mellitus due to complete glucokinase deficiency. *New England Journal of Medicine*, **344**, 1588–92.

Ozanne, S. E. and Hales, C. N. (1998). Thrifty yes, genetic no. *Diabetologia*, **41**, 485–7.

Ozanne, S. E. (2001). Metabolic programming in animals. *British Medical Bulletin*, **60**, 143–52.

Pan, X. R., Li, G. W., Hu, Y. H. *et al.* (1997). Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance. The Da Qing IGT and Diabetes Study. *Diabetes Care*, **20**, 537–44.

Parker, A., Meyer, J., Lewitzky, S. *et al.* (2001). A gene conferring susceptibility to type 2 diabetes in conjunction with obesity is located on chromosome 18p11. *Diabetes*, **50**, 675–80.

Patti, M. E., Butte, A. J., Crunkhorn, S. *et al.* (2003). Coordinated reduction of genes of oxidative metabolism in humans with insulin resistance and diabetes: potential role of PGC1 and NRF1. *Proceedings of the National Academy of Sciences USA*, **100**, 8466–71.

Pearson, E. R., Starkey, B. J., Powell, R. J. *et al.* (2003). Genetic cause of hyperglycaemia and response to treatment in diabetes. *Lancet*, **362**, 1275–81.

Permutt, M. A., Wasson, J. C., Suarez, B. K. *et al.* (2001). A genome scan for type 2 diabetes susceptibility loci in a genetically isolated population. *Diabetes*, **50**, 681–5.

Perry, R. and Baron, A. (1999). Impaired glucose tolerance. Why is it not a disease? *Diabetes Care*, **22**, 920–4.

Pharoah, P. D., Antoniou, A., Bobrow, M. *et al.* (2002). Polygenic susceptibility to breast cancer and implications for prevention. *Nature Genetics*, **31**, 33–6.

Phillips, D. I. W. (1993). Twin studies in medical research: can they tell us whether diseases are genetically determined? *Lancet*, **341**, 1008–10.

Phillips, D. I. W. (1996). Insulin resistance as a programmed response to fetal undernutrition. *Diabetologia*, **39**, 1119–22.

Poulsen, P., Vaag, A. A., Kyvik, K. O., Møller Jensen, D. and Beck-Nielsen, H. (1997). Low birth weight is associated with NIDDM in discordant monozygotic and dizygotic twin pairs. *Diabetologia*, **40**, 439–46.

Poulsen, P., Kyvik, K. O., Vaag, A. and Beck-Nielsen, H. (1999). Heritability of Type II (non-insulin-dependent) diabetes mellitus and abnormal glucose tolerance – a population-based twin study. *Diabetologia*, **42**, 139–45.

Proks, P., Reimann, F., Green, N., Gribble, F. and Ashcroft, F. (2002). Sulfonylurea stimulation of insulin secretion. *Diabetes*, **51**, Suppl 3, S368–S376.

Raeder, H., Johansson, S., Holm, P. I. *et al.* (2006). Mutations in the *CEL* VNTR cause a syndrome of diabetes and pancreatic exocrine dysfunction. *Nature Genetics*, **38**, 54–62.

Ravelli, A. C. J., van der Meulen, J. H. P., Michels, R. P. J. *et al.* (1998). Glucose tolerance in adults after prenatal exposure to famine. *Lancet*, **351**, 173–7.

Saltiel, A. R. and Kahn, C. R. (2002). Insulin signalling and the regulation of glucose and lipid metabolism. *Nature*, **414**, 799–806.

Silander, K., Mohlke, K. L., Scott, L. J. *et al.* (2004). Genetic variation near the Hepatocyte Nuclear Factor-4 gene predicts susceptibility to Type 2 Diabetes. *Diabetes*, **53**, 1141–9.

Sotos, J. G., Reinhoff, H. Y., Block, G. D. *et al.* (2000). Will genetics revolutionize medicine? *New England Journal of Medicine*, **343**, 1496–8.

Stanner, S. A., Bulmer, K., Andrès, C. *et al.* (1997). Does malnutrition in utero determine diabetes and coronary heart disease in adulthood? Results from the Leningrad Siege Study, a cross sectional study. *British Medical Journal*, **315**, 1342–9.

Stoffers, D. A., Ferrer, J., Clarke, W. L. and Habener, J. F. (1997). Early-onset type-II diabetes mellitus (MODY4) linked to IPF1. *Nature Genetics*, **17**, 138–9.

Stone, L. M., Kahn, S. E., Fujimoto, W. Y., Deeb, S. S. and Porte, D. (1996). A variation at position − 30 of the β-cell glucokinase gene promoter is associated with reduced β-cell function in middle-aged Japanese-American men. *Diabetes*, **45**, 422–8.

Tattersall, R. B. and Mansell, P. I. (1991). Maturity onset-type diabetes of the young (MODY): one condition or many? *Diabetic Medicine*, **8**, 402–10.

The Expert Committee on the Diagnosis and Classification of Diabetes Mellitus (1997). Report of the expert committee on the diagnosis and classification of diabetes mellitus. *Diabetes Care*, **20**, 1183–97.

Triggs-Raine, B. L., Kirkpatrick, R. D., Kelly, S. L. *et al.* (2002). HNF-1alpha G319S, a transactivation-deficient mutant, is associated with altered dynamics of diabetes onset in an Oji-Cree community. *Proceedings of the National Academy of Sciences USA*, **99**, 4614–19.

Tuomilehto, J., Lindstrom, J., Eriksson, J. G. *et al.*, Finnish Diabetes Prevention Study Group (2001). Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *New England Journal of Medicine*, **344**, 1343–50.

UK Prospective Diabetes Study (UKPDS) Group (1998). Intensive blood-glucose control with sulphonylureas or

insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). *Lancet*, **352**, 837–53.

Vionnet, N., Hani, E. H., Dupont, S. *et al.* (2000). Genome-wide search for type 2 diabetes-susceptibility genes in French Whites: evidence for a novel susceptibility locus for early-onset diabetes on chromosome 3q27-qter and independent replication of a type 2-diabetes locus on chromosome 1q21-q24. *American Journal of Human Genetics*, **67**, 1470–80.

Virkamaki, A., Korsheninnikova, E., Seppala-Lindroos, A. *et al.* (2001). Intramyocellular lipid is associated with resistance to in vivo insulin actions on glucose uptake, antilipolysis, and early insulin signaling pathways in human skeletal muscle. *Diabetes*, **50**, 2337–43.

Watanabe, R. M., Valle, T., Hauser, E. R. *et al.* (1999). Familiality of quantitative metabolic traits in Finnish families with non-insulin-dependent diabetes mellitus. Finland–United States investigation of non-insulin-dependent diabetes mellitus genetics (FUSION) study investigators. *Human Heredity*, **49**, 159–68.

Weedon, M. N., Frayling, T. M., Shields, B. *et al.* (2005). Genetic regulation of birth weight and fasting glucose by a common polymorphism in the islet cell promoter of the glucokinase gene. *Diabetes*, **54**, 576–81.

Weedon, M. N., Owen, K. R., Shields, B. *et al.* (2004). Common variants of the HNF4alpha P2 promoter are associated with type 2 diabetes in the UK population. *Diabetes*, **53**, 3002–6.

Weedon, M. N., Schwarz, P. E. H., Horikawa, Y. *et al.* (2003). Meta-analysis confirms a role for calpain-10 variation in type 2 diabetes susceptibility. *American Journal of Human Genetics*, **73**, 1208–12.

Wiltshire, S., Hattersley, A. T., Hitman, G. A. *et al.* (2001). A genome-wide scan for loci predisposing to type 2 diabetes in a UK population (The Diabetes (UK) Warren 2 Repository): analysis of 573 pedigrees provides independent replication of a susceptibility locus on chromosome 1q. *American Journal of Human Genetics*, **69**, 553–69.

World Health Organization Study Group (1985). Diabetes Mellitus: Report of a WHO Study Group. *World Health Organ Technical Report Series*, **727**, 1–113.

Yamagata, K., Furuta, H., Oda, N. *et al.* (1996*a*). Mutations in the hepatocyte nuclear factor-4α gene in maturity-onset diabetes of the young (MODY1). *Nature*, **384**, 458–60.

Yamagata, K., Oda, N., Kaisaki, P. J. *et al.* (1996*b*). Mutations in the hepatocyte nuclear factor-1α gene in maturity-onset diabetes of the young (MODY3). *Nature*, **384**, 455–8.

Yang, Q., Khoury, M. J., Botto, L., Friedman, J. M. and Flanders, W. D. (2003). Improving the prediction of complex diseases by testing for multiple disease-susceptibility genes. *American Journal of Human Genetics*, **72**, 636–49.

Yki-Jarvinen, H. (2002). Ectopic fat accumulation: an important cause of insulin resistance in humans. *Journal of the Royal Society of Medicine*, **95** Suppl 42, 39–45.

Zimmet, P. (2000). Globalization, coca-colonization and the chronic disease epidemic: can the doomsday scenario be averted? *Journal of Internal Medicine*, **247**, 301–10.

Zimmet, P., Alberti, K. G. M. M. and Shaw, J. (2001). Global and societal implications of the diabetes epidemic. *Nature*, **414**, 782–7.

Zung, A., Glaser, B., Nimri, R. and Zadik, Z. (2004). Glibenclamide treatment in permanent neonatal diabetes mellitus due to an activating mutation in Kir6.2. *Journal of Clinical Endocrinology and Metabolism*, **89**, 5504–7.

24

Genetics of coronary heart disease

Rossi Naoumova, Stuart A. Cook, Paul Cook and Timothy J. Aitman

Coronary heart disease (CHD) is the single commonest cause of death in the developed world. One in four men and one in six women die from CHD. In the UK, around 15% of these deaths occur under the age of 65, and 35% under the age of 75 (www.heartstats.org). CHD frequency varies between populations, with the highest age-adjusted rates of 600–1000 deaths per 100 000 found in countries of the former Soviet Union and the lowest at around 60 per 100 000 in Japan (World Health Organization 2002, www.who.ch/). Age-adjusted rates in the UK and USA are around 200–300 per 100 000 of population. Age-adjusted CHD prevalence in the UK and USA has fallen by around 40% in the past 30 years, although this reflects more a postponement in age of CHD-related death by about 10 years rather than an absolute reduction in numbers of deaths (Fuster, 1999). CHD is predicted to remain the commonest single cause of death in developed countries over the next 20 years and will increase in frequency to become the commonest cause of disease-related disability in both developed and developing countries by the year 2020 (Murray and Lopez, 1997).

Genetic and environmental contributions to CHD pathogenesis

The significant changes in CHD incidence and mortality over the past 20 years can be attributed at least in part to variation in known environmental risk factors (Table 24.1). Frequency of cigarette smoking, efficacy of primary and in-hospital coronary care, treatment of plasma lipids, and the prevalence of obesity and type 2 diabetes (T2DM) have changed significantly over the past 20 years and also vary markedly between populations. These environmental factors explain, at least in part, the widely differing CHD prevalence and trends in CHD prevalence observed between populations over the past two to three decades.

Whilst environmental influences account for a significant part of individual and population susceptibility to CHD, the longevity of some individuals with multiple risk factors (e.g. smoking, hypercholesterolemia, hypertension), and the consistent and highly significant risk conferred by a positive family history of CHD, point to the importance of genetic factors in CHD pathogenesis. Genes which control plasma lipids and lipoproteins are amongst the clearest examples of genetic risk factors for CHD (Table 24.1).

In this chapter, we discuss first the evidence from family studies that genetic influences are important in susceptibility to CHD. Secondly, we review the genetics of lipid risk factors, including the genetic basis of monogenic and complex dyslipidaemias, and their impact on CHD susceptibility. Thirdly, we review genetic studies of human CHD and its risk factors, including consideration of the role played by the metabolic syndrome, whose prevalence is increasing rapidly

Table 24.1. Known risk factors for coronary heart disease

Environmental (modifiable)	Mainly genetic (Primary cause not modifiable)	Combined genetic and environmental
Cigarette smoking	Lp(a)	Obesity
Physical inactivity	Monogenic dyslipidemias	Raised LDL cholesterol, triglycerides; reduced HDL cholesterol
Dietary fat	Ethnic origin	Complex dyslipidemias (e.g. familial combined hyperlipidemia)
Lipid lowering therapies	Family history	Insulin resistance, metabolic syndrome and type 2 diabetes
		Small dense low density lipoprotein
		Homocysteine
		Fibrinogen, PAI-1
		C-reactive protein

LDL: low density lipoprotein; HDL: high density lipoprotein; PAI-1: plasminogen activator inhibitor-1.

and whose contribution to CHD prevalence and susceptibility is increasingly important. Finally, we discuss data from rodent studies which have thrown light on the genes, regulatory pathways and molecular mechanisms underlying development of dyslipidemia and atherosclerosis.

The understanding of atherosclerosis cell biology and pathophysiology has advanced dramatically over the past 10–15 years. These areas have been reviewed extensively elsewhere (Glass and Witztum, 2001; Lusis, 2000; 2003), and will not be discussed here unless of direct relevance to the genetic basis of CHD.

Evidence for the genetic basis of CHD

Family history of the disease is one of the strongest and most consistent risk factors for CHD. Relative risk estimates of between 1.5 and 5 have been derived from a range of family and population studies (Hawe *et al.*, 2003; Lusis, 2003). At least a part of the risk of family history is conferred by the monogenic dyslipidaemias, including familial hypercholesterolemia (FH) (Goldstein *et al.*, 1973), which carry a high penetrance of CHD. However, since the frequency of the FH heterozygous state, and of other monogenic dyslipidaemias, is only at most around 1 in 500, these cannot account for more than a small proportion of all cases.

Although the risk conferred by family history is consistent, such family studies cannot distinguish reliably shared environmental from inherited influences. More informative are twin and adoption studies, which confirm definitively the part played by inherited factors. A 26-year follow-up of 21 000 Swedish twins showed a relative hazard of 8.1 for coronary disease in a male monozygotic twin dying of CHD before age 55 years compared to monozygotic twins not dying of CHD, and a relative hazard of 3.8 for dizygotic twins. Among female twins, the relative hazard for the monozygotic twin dying under age 65 years was 15.0 compared to 2.6 for a dizygotic twin. Heritability after 36 years follow-up was estimated to be 0.57 in male twins and 0.38 in female twins (Zdravkovic *et al.*, 2002). Such a high heritability is supported by the adoption study of Sorensen *et al.* (Sorensen *et al.*, 1988) who found a relative risk of death in adoptees of 4.5 for death from vascular disease before age 50 years of a biological parent, compared to a relative risk of 3.0 for death of an adoptive parent. These results and similar data from other studies provide clear evidence for the existence of CHD susceptibility genes and provide the justification for the large number of studies seeking their identification.

Genetics of monogenic and complex dyslipidemias

Disorders of lipid metabolism, both monogenic and more genetically complex, are powerful risk factors for CHD. The major circulating lipids, cholesterol, phospholipids and triglycerides, are carried in the blood by several lipoprotein particles serving the physiological task of transporting dietary and endogenously produced lipids to peripheral tissues. The major pathways of lipoprotein metabolism in humans are illustrated in (Figure 24.1). Mutations in many of the lipoproteins, receptors and enzymes shown in the figure have been described in humans or engineered in rodents. Studies of these mutations have yielded the molecular basis of a wide range of dyslipidemias and in many cases elucidated the mechanism by which dysfunction of these proteins predispose to, or protect against, atherosclerosis and CHD.

Monogenic hypercholesterolemias

Important insights into CHD pathogenesis have come from the study of monogenic hypercholesterolemias, and have led to the development of the most widely used class of lipid-lowering drugs, the HMG-CoA-reductase inhibitors (statins). Among the many known risk factors for atherosclerosis, very high low density lipoprotein (LDL) cholesterol levels are probably unique in their ability to lead to the development of premature atherosclerosis in humans in the absence of other additional risk factors. Table 24.2 summarizes the main characteristics of the known Mendelian disorders of severe hypercholesterolemia, all of which are associated with premature CHD.

Autosomal dominant hypercholesterolemias

Familial hypercholesterolemia (FH) was the first genetic disease of lipid metabolism to be clinically and genetically characterized and it represents the most common and most severe form of Mendelian hypercholesterolemia (Goldstein *et al.*, 2001). The disease is caused by mutations in the LDL receptor (*LDLR*) and so far nearly 900 mutations have been identified (www.ucl.ac.uk/fh). In FH homozygotes serum levels of LDL cholesterol are very high from birth, irrespective of lifestyle, and response to treatment depends to a large extent on the type of *LDLR* mutation and residual LDLR activity. FH homozygotes develop severe atherosclerosis in early childhood, initially in the aortic root, causing supravalvular aortic stenosis and then involving the coronary ostia and arteries. If treatment to reduce LDL cholesterol is not initiated in early childhood or is inadequate, FH homozygotes die of myocardial infarction very early in life, often in the first decade (Goldstein *et al.*, 2001; Thompson, 1999). In patients with heterozygous FH, serum LDL cholesterol levels are lower than in homozygotes and clinical prognosis depends much more on other genetic and environmental factors (Hill *et al.*, 1991; Pimstone *et al.*, 1998). If untreated, more than 50% of male and about 15% of female FH heterozygotes die before 60 years of age of CHD (Marks *et al.*, 2003; Slack, 1969).

A clinically similar syndrome to FH is familial defective apolipoproteinB-100 (FDB), which results from a mutation in the LDLR binding domain of the apolipoproteinB (*ApoB*) gene (Soria 1989). FDB homozygotes have much lower levels of serum LDL cholesterol than FH homozygotes and the development of atherosclerotic complications is delayed (Myant, 1993). LDL cholesterol levels are lower in FDB than in FH due to lack of accumulation of cholesterol rich remnant particles, which utilize apolipoprotein E as a ligand for the LDLR (Figure 24.1) (Schaefer *et al.*, 1997).

A new form of autosomal dominant hypercholesterolemia, caused by mutations in the *PCSK9* gene has been recently described (Abifadel *et al.*, 2003). The gene, residing on chromosome 1, encodes neutral apoptosis-regulated convertase 1 (NARC-1), a member of the proteinase K family of substances (Seidah *et al.*, 2003). Some missense mutations in the *PCSK9* gene have been associated with unusually severe hypercholesterolemia and

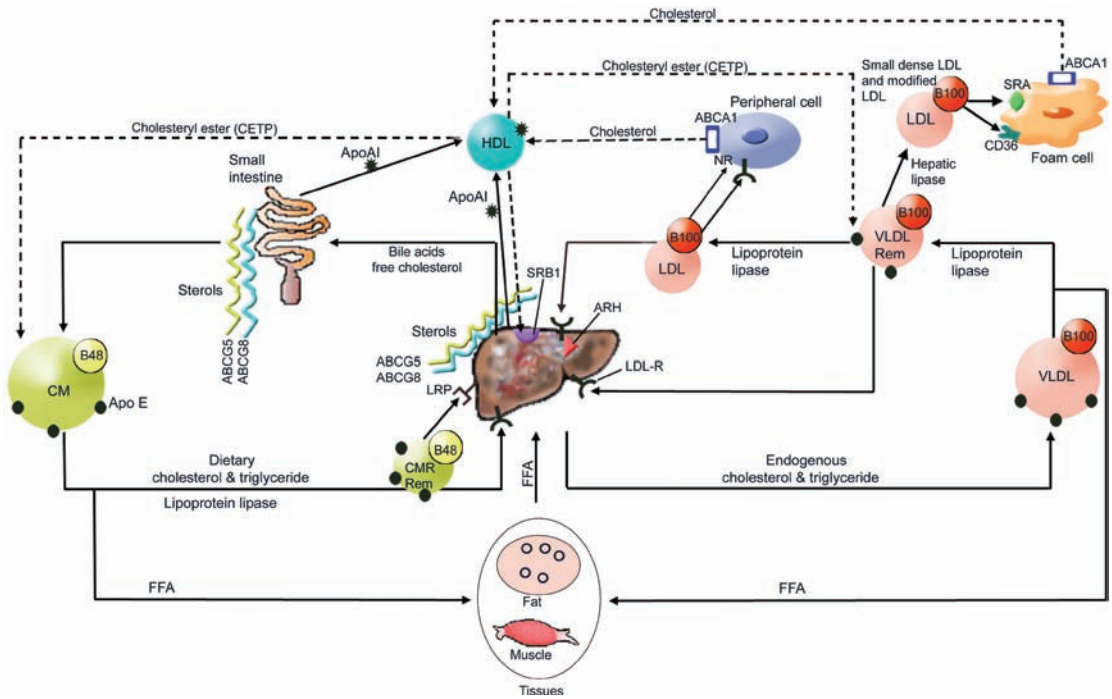

Figure 24.1 Outline of lipoprotein metabolism in humans. Chylomicrons (CM) transport dietary lipids, while very low density lipoproteins (VLDL), low density lipoproteins (LDL) and the high density (HDL) lipoproteins transport endogenously synthesized lipids. Apolipoprotein (apo)B is the major structural protein of CM, VLDL and LDL, and apoA-I of HDL.

Dietary lipids are packed with apolipoproteins in the small intestine to form CM. In the circulation the core triglycerides of CM are hydrolyzed and form CM remnants. VLDL particles secreted into the plasma by the liver are triglyceride rich and undergo lipolysis to form cholesterol-rich VLDL remnants. Around half of the VLDL remnants are removed by the liver via the LDL receptor (LDLR), for which both apoE and apoB are ligands. The remainder of VLDL are further converted into LDL particles, which are cleared from the circulation by the LDLR or by the non-receptor mediated pathway. Nascent HDL during its transformation into HDL is directly involved in reverse cholesterol transport by acquiring lipids, providing an antiatherogenic function of HDL. Passage of cholesterol across the cell membrane depends on an ATP-binding cassette transporter (ABCA1), which is a cholesterol efflux regulatory protein. A significant part of the cholesterol transported to different tissues is normally returned to the liver for elimination in bile as bile acids and free cholesterol, a process for which the two ABC half-transporters: ABCG5 and ABCG8 are required. Reviewed in more detail elsewhere (Durrington, 2003; Packard and Shepherd 1999). Adapted from Shoulders *et al.*, 2004.

CETP = cholesterol ester transfer protein; NR = non-receptor mediated transport; SRA = scavenger receptor A; SRB1 = scavenger receptor B1; FFA= free fatty acids; ARH = autosomal recessive hypercholesterolemia, a putative LDLR adaptor protein; LRP = LDL receptor-like protein.

very early CHD (Naoumova *et al.*, 2005) due, at least in part, to overproduction of ApoB (Sun *et al.*, 2005). Elucidation of the substrates and function of *PCSK9* gene product and the mechanisms involved in the development of severe hypercholesterolemia will provide new insights into the metabolism of atherogenic lipoproteins and possibly novel therapeutic targets.

Table 24.2. Monogenic disorders that cause hypercholesterolemia

Disease	Defective gene	Prevalence	Plasma LDL-C	Plasma TG	Premature CHD risk	Major metabolic defect	Reference
Autosomal dominant forms							
FH (''classic'')	*LDLR*					↓ LDL clearance	(Goldstein *et al.*, 2001)
Heterozygous FH		1:500	+++	− to +	+++		
Homozygous FH		$1:1\times10^6$	++++++	− to +	++++++		
FDB	*ApoB*					↓ LDL clearance	(Myant, 1993; Soria *et al.*, 1989)
Heterozygous FDB		1:1000	++	−	++		
Homozygous FDB		$1:4\times10^6$	++++	−	+++		
FH_3	*PCSK9 (NARC-1)*					Unknown	(Abifadel *et al.*, 2003; Naoumova *et al.*, 2005; Varret *et al.*, 1999)
Heterozygous FH_3		3 families reported	++++	− to +	++++		
Homozygous FH_3		?	?	?	?		
Autosomal recessive forms							
Autosomal recessive hypercholesterolemia (ARH)	*ARH*	~50 patients reported	+++++	− to +	+++++	↓ LDL clearance	(Arca *et al.*, 2002; Garcia *et al.*, 2001; Naoumova *et al.*, 2004)
Sitosterolemia	*ABCG5* or *ABCG8*	~50 patients reported	− to +++++	−	+++++	↑ sterol absorption ↓ sterol secretion in bile	(Berge *et al.*, 2000; Salen *et al.*, 1992)
Cholesterol 7α-hydroxylase deficiency	*CYP7A1*	3 siblings reported	++	− to ++	Unknown	↓ Cholesterol excretion ↓ LDL clearance ?	(Pullinger *et al.*, 2002)

FH: familial hypercholesterolemia; FDB: familial defective apolipoprotein B; LDL: low density lipoprotein.

Autosomal recessive hypercholesterolemias

The recent characterization of a rare gene defect causing autosomal recessive hypercholesterolemia (ARH) (Garcia *et al.*, 2001) has provided new information into the underlying mechanism of clathrin-mediated internalization of the LDLR. Mutations in *ARH*, which encodes a novel adaptor protein, prevent normal internalization of the LDLR by cultured lymphocytes and monocyte-derived macrophages, but not by skin fibroblasts (Eden *et al.*, 2002). The clinical phenotype of ARH is similar to that of classical homozygous FH caused by defects in the *LDLR* gene, but is more variable and generally less severe (Naoumova *et al.*, 2004).

Sitosterolemia is a rare but highly atherogenic disorder, characterised by the accumulation of both animal and plant sterols in the blood and tissues and development of aortic stenosis and premature CHD (Salen *et al.*, 1992). Sitosterolemia results from a defect in sterol efflux from cells and is caused by mutations in either of two adjacent genes that encode ABC half-transporters, *ABCG5* and *ABCG8*, expressed almost exclusively in hepatocytes and enterocytes (Figure 24.1) (Berge *et al.*, 2000; Lee *et al.*, 2001).

A hypercholesterolemic phenotype similar to that in heterozygous FH has also been reported in siblings with homozygous mutations in the gene for 7-α-hydroxylase, the first enzyme in the pathway of bile acid synthesis in the liver (Pullinger *et al.*, 2002).

Monogenic disorders causing impaired reverse cholesterol transport

Low HDL cholesterol is a common finding in patients with premature CHD and usually precedes clinically apparent CHD (Goldbourt *et al.*, 1997). Low HDL concentrations are most frequently associated with the components of the metabolic syndrome X and are rarely due to single-gene disorders (Durrington, 2003). The rare Tangier disease, characterized by very low or even immeasurable high density lipoprotein (HDL) cholesterol concentrations and reduced LDL cholesteryl ester, is due to mutations in the ATP-binding cassette (*ABCA1*) transporter gene encoding cholesterol-efflux regulatory protein (Oram, 2002) (see Figure 24.1). Patients with familial HDL deficiency, who are heterozygotes for *ABCA1* mutations, have cholesterol concentrations between 0.4 and 0.9 mmol/l and are at increased risk of premature CHD (Marcil *et al.*, 1999). Mutations in the *ApoAI* gene, leading to under-expression of the apoprotein cause low HDL and premature CHD (Tall, 1998).

Conversely, cholesterol ester transfer protein deficiency, which is a common cause for high HDL concentrations (>2.5 mmol/l) in Japan is thought to be a state of impaired reverse cholesterol transport, which may lead to premature CHD (Nagano *et al.*, 2002). Thus, HDL cholesterol levels, efflux of cellular lipids as well as the kinetics of HDL metabolism are important determinants of the antiatherogenic effect of HDL (Tall, 1998; von Eckardstein, 2001).

Oligo/polygenic (complex) genetic hyperlipidemias

Familial combined hyperlipidemia (FCHL)

Combined hyperlipidemia, characterized by increased serum triglyceride and/or cholesterol and decreased HDL cholesterol is the most common disorder of lipid metabolism affecting 1–2% of individuals in Western societies.

The term familial combined hyperlipidemia (FCHL) was coined by Goldstein *et al.* to describe a pattern of lipid abnormalities (Goldstein *et al.*, 1973) in survivors of myocardial infarction who had raised blood cholesterol and triglyceride levels. Originally, FCHL was described as a dominant disorder with a primary effect on blood triglyceride levels, a secondary effect on cholesterol levels, and with incomplete penetrance until the third decade (Goldstein *et al.*, 1973). However subsequent segregation analyses (Austin, 1992;

Cullen *et al.*, 1994) and three genome-wide studies (Aouizerat *et al.*, 1999; Naoumova *et al.*, 2003; Pajukanta *et al.*, 1999) have suggested a more complex inheritance pattern (Shoulders *et al.*, 2004).

The molecular basis of FCHL is still unknown. The evidence for major genes acting on apoB, insulin and triglyceride levels in FCHL is consistent with the metabolic finding of increased production of VLDL and *apoB100* (Kissebah *et al.*, 1984; Venkatesan *et al.*, 1993), defective catabolism of VLDL and chylomicrons (Babirak *et al.*, 1992; Cabezas *et al.*, 1993) and increased production of *apoCIII* and insulin resistance (Aitman *et al.*, 1997*a*; Dallinga-Thie *et al.*, 1996; Ito *et al.*, 1990).

Candidate gene approaches to identify the primary metabolic defects underlying FCHL have been largely unrewarding. Genome-wide studies in FCHL-implicated regions on chromosome 1q, 6q, 8p and 11p (Aouizerat *et al.*,1999; Naoumova *et al.*, 2003; Pajukanta *et al.*, 1999). *Apo A5*, a newly discovered gene involved in lipid metabolism (Pennacchio *et al.*, 2001), was found to contribute to the transmission of FCHL in a proportion of British FCHL families (Eichenbaum-Voline *et al.*, 2004).

Type III hyperlipoproteinemia (dysbetalipoproteinemia or broad β disease)

Polymorphism of the *apoE* gene has been detected in all populations. The most common allele in Europe is *apoE3* (frequency 0.75) followed by *apoE4* (0.13) and *apoE2* (0.1) (Motulsky *et al.*, 2002). *ApoE3* binds with high affinity to the LDLR, whereas *apoE2* has much reduced binding. About 1% of the population are homozygous for *apoE2* but only 1% of these individuals will have the additional genetic or acquired factors (obesity, diabetes, hypothyroidism, FCHL, alcohol abuse) that lead to type III hyperlipidaemia. This rare disorder, which results from the presence in the circulation of large amounts of CM remnants and VLDL remnants (collectively termed β VLDL) is characterized by markedly raised serum cholesterol and triglyceride, palmar striae and tuberoeruptive xanthomata. If left untreated patients have strikingly increased risk for premature CHD and peripheral vascular disease (Mahley *et al.*, 2001).

Genetics of human coronary heart disease

Because the classic CHD risk factors such as serum cholesterol, smoking, hypertension and diabetes mellitus are incomplete predictors of CHD, additional risk factors based on molecular genetics have been intensively investigated. The two main methods of investigation to date are genome-wide linkage studies and case-control association studies. Whole-genome association studies using ultra high-throughput genotyping have yet to report, although are expected to do so within the next 2–3 years. Most studies to date have focused either on identifying genes that predispose to the traditional risk factors, or on direct linkage and association to the phenotype of CHD per se. We shall focus here on genetic studies of the CHD phenotype itself. The genetics of lipid risk factors has been reviewed above and the genetics of other CHD risk factors including diabetes and hypertension are reviewed in other chapters in this book (Chapters 23, 25).

Genome-wide studies for coronary heart disease

Five independent, genome-wide screens for CHD have been published on Finnish, Mauritian, Australian and European families with various gene localizations reported (Francke *et al.*, 2001; Harrap *et al.*, 2002; Broeckel *et al.*, 2002; Helgadottir *et al.*, 2004; BHF Family Heart Study Research Group, 2005). Most studies have been inconclusive with no LOD score achieving the threshold for genome-wide statistical significance although one study (involving Western European families) localised a new CHD QTL on chromosome 14 with a significant LOD score of 3.9 (Broeckel *et al.*, 2002). This region does not overlap

with existing QTLs for other CHD risk factors or intermediate phenotypes, raising the possibility that, whilst the role of existing risk factors in this QTL are not excluded, this QTL may act through a presently unknown mechanism.

The heritability of known risk factors that were included in this linkage study (including type 2 diabetes mellitus and hypertension) was high, suggesting that these intermediate phenotypes are genetically determined to a moderately high extent. However, whereas age, gender, diabetes and hypertension contributed significantly to the myocardial infarction phenotype, high cholesterol levels did not. Since many of the individuals in the study were on lipid-lowering therapy, this suggests that hypercholesterolemia now appears to be a less significant risk factor for myocardial infarction than diabetes and hypertension.

The genome scan in Icelandic subjects localized a susceptibility gene for both myocardial infarction and stroke to chromosome 13 and identified the likely gene as 5-lipoxygenase activating protein (*FLAP*) (Helgadottir *et al.*, 2004). The *FLAP* gene product converts 5-lipoxygenase to leukotriene B4, a powerful inflammatory cytokine, emphasising the importance of inflammation in CHD pathogenesis, and raising the possibility of new therapeutic strategies targeting this and other genes in inflammatory pathways.

A meta-analysis was applied to four of the CHD genome-wide studies (Chiodini and Lewis, 2003). The genetic region 3q26–27 showed the strongest evidence for linkage ($P = 0.0001$), which is one of the regions of linkage found in the Mauritian population (Francke *et al.*, 2001). This result is of interest, and the subject of further investigation, given the existence of several candidate genes involved in the homeostasis of glucose and lipid metabolism in this chromosomal region and the important association between CHD and the metabolic syndrome.

Informative mutations have also been found by utilizing families that demonstrate Mendelian pattern of inheritance for CHD. Wang and colleagues carried out a genome scan in a single family demonstrating a rare autosomal dominant pattern of inheritance for CHD (Wang *et al.*, 2003). The cause of the CHD phenotype in this kindred was proposed to be a deletion of seven amino acids in a member of the myocyte enhancer factor-2 family of transcription factors (*MEF2A*). The *MEF2A* protein showed strong expression in the endothelium of coronary arteries, but how this deletion gives rise to the CHD phenotype remains to be determined. Since the original observations of Wang *et al.*, the association of *MEF2A* with CHD has both been challenged and replicated (Altshuler and Hirschhorn, 2005; Topol, 2005). Although debate therefore exists on the definitive nature of these findings, the data imply that genes in the *MEF2A* signalling pathway, as well as genes regulating endothelial development and function, may be important in the pathophysiology of CHD.

Candidate gene studies for coronary heart disease

When selecting candidate genes, investigators must choose from a very large number of potential factors that may contribute to the CHD phenotype. These are generally deduced from knowledge of the pathophysiological pathways involved. As a consequence, genes involved in lipid metabolism, the renin-angiotensin-aldosterone system, thrombosis, insulin resistance, inflammation and endothelial function represent potential candidates for CHD. A full survey of candidate genes is beyond the scope of this article and has been reviewed elsewhere (Navarro-Lopez, 2002). Therefore only a small number of illustrative examples of candidate genes shown to be associated with CHD will be given.

One of the best-studied genetic variants is the insertion/deletion (I/D) polymorphism of the angiotensin converting enzyme (*ACE*) gene. ACE is predominantly found in the vascular endothelium of endothelial cells and catalyzes the generation of angiotensin II from angiotensin I and the degradation of bradykinin. The ACE polymorphism is characterized by the presence or absence of an

Alu sequence of 287 base pairs in intron 16. It gives rise to three different genotypes: II/ID/DD with carriers of the D allele having an increase in plasma, cardiac and tissue ACE activity. In 1992, the association of the DD genotype with an increased risk of myocardial infarction was reported (Cambien *et al.*, 1992). As with many candidate gene studies, this association has been variably reproduced. In particular a large prospective population study performed by Lindpaintner and colleagues did not confirm the association (Lindpaintner *et al.*, 1995).

Factor VII participates with tissue factor at the onset of the coagulation cascade. A reduction in factor VII levels reduces the propensity to thrombosis. Two polymorphisms have been implicated as having a protective role against myocardial infarction. The FVII R353Q polymorphism, with substitution of arginine (R) by glutamine (Q) in codon 353 of exon 8, and the 5′ F7 polymorphism, a decanucleotide insertion at position -323 of the 5′ region of the promoter, with two alleles, A1 (without the insertion) and A2 (with the insertion). People who are homozygous for the Q or A2 genotype have plasma levels of activated factor VII which are 70% lower than in R or A1 homozygotes. In one Italian study the Q or A2 allele was associated with a reduction in the risk of myocardial infarction (Iacoviello *et al.*, 1998). Another Italian study replicated these results and showed that these polymorphisms were major determinants of factor VII levels (Girelli *et al.*, 2000). Failure to replicate the association with CAD in other investigations has been ascribed to geographical variation in the frequency of these polymorphisms, implying that these polymorphisms are not themselves the functional variants underlying altered factor VII or CHD susceptibility, although other closely linked polymorphisms might underlie these significant phenotypes (Doggen *et al.*, 1998; Lane *et al.*, 1996).

Three further candidate gene studies merit particular attention. Arising from a very high throughput screen of over 92 000 single nucleotide polymorphisms (SNPs) in a study of over 3000 cases and controls, Ozaki and colleagues showed strong association ($P < 10^{-5}$) between susceptibility to myocardial infarction and SNPs in the lymphotoxin-α gene (Ozaki *et al.*, 2002). Subsequently the same research group showed association of myocardial infarction with another gene on the same pathway, galectin-2, which plays a part in the regulation of lymphotoxin-α secretion (Ozaki *et al.*, 2004). Most recently Helgadottir *et al.* (2006) demonstrated strong association between leukotriene A4 hydrolase and myocardial infarction in African-Americans, and replication in European-Americans, though with less signficance, again implicating inflammatory pathways in the pathogenesis of myocardial infarction, but also suggesting interaction between inflammatory and other pathways.

Insulin resistance has a central role in the development of hypertension, dyslipidemia, and T2DM (Figure 24.2) and, with the current epidemic of obesity and T2DM, it is a contributory factor in an increasing proportion of cases of premature

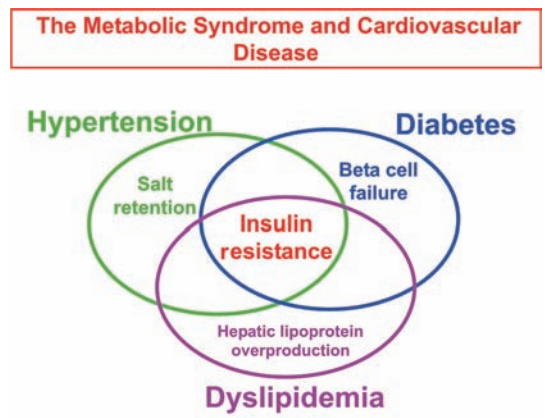

Figure 24.2 Insulin resistance and the development of major cardiovascular risk factors. Insulin resistance alone has subtle effects that may be detected biochemically, but when combined with pancreatic beta cell failure, renal salt retention or hepatic overproduction of lipoproteins, leads to clinically overt diabetes, hypertension and dyslipidemia. These are amongst the strongest and most prevalent risk factors for coronary heart disease.

coronary heart disease, estimated to be as high as 60% in some populations (McKeigue *et al.*, 1993; Reaven, 1988). Since the original description of the Metabolic Syndrome, its component features have expanded to include microalbuminuria, central obesity, raised levels of plasminogen activator inhibitor-1 (PAI-1) and uric acid. Genetic associations have been shown in case-control studies, although many of these still require replication and mechanistic explanation. The most consistent of these associations is that of the peroxisome proliferator activated receptor (PPAR)-γ polymorphism Pro12Ala with type 2 diabetes mellitus and insulin sensitivity. By analyzing over 3000 individuals, the authors found a modest (1.25-fold) but significant ($P = 0.002$) increase in diabetes risk associated with the more common proline allele (~85% frequency). Because the risk allele occurs at such high frequency, its modest effect translates into a large population attributable risk, influencing as much as 25% of type 2 diabetes in the general population (Altshuler *et al.*, 2000).

The possibility that Cd36 deficiency in humans might be a cause of insulin resistance arose initially from rodent studies (Aitman *et al.*, 1997*b*; 1999; Pravenec *et al.*, 2001). Cd36 is a multi-functional protein and belongs to the scavenger receptor family typified by SrB1 and mediates the specific uptake of long chain-free fatty acids by adipocytes and muscle cells. Cd36 is also a high affinity receptor for oxidized low density lipoproteins, a role which has been suggested to play a part in macrophage foam cell formation and atherosclerosis (Tontonoz *et al.*, 1998). Varying rates of prevalence of CD36-deficiency in different ethnic groups have been documented but are as high as 1% in black Africans and African-Caribbeans. Japanese subjects with CD36-deficiency have increased levels of triglycerides, increased fasting plasma glucose, hypertension and reduced insulin action (Miyaoka *et al.*, 2001). Importantly, diabetes and insulin resistance are more common in CD36-deficient Japanese than in the wider Japanese population (Furuhashi *et al.*, 2003). Whether heterozygous or homozygous mutations for human Cd36-deficiency predispose or protect against CHD has yet to be determined.

Finally, mention should be made of two novel genetic mechanisms that may play a part in CHD pathogenesis. Samani and colleagues showed evidence for telomere shortening in subjects with atherosclerosis, showing a possible association of accelerated chromosomal ageing with CHD (Samani *et al.*, 2001). Other studies have implicated structural rearrangements in the genome, such as gene duplication and deletion, as a cause of common human diseases such as HIV susceptibility and glomerulonephritis (Gonzalez *et al.*, 2005; Aitman *et al.*, 2006). Although specific studies to test the hypothesis that structural rearrangements predispose to CHD have as yet not been carried out, such studies, if positive, have the potential to elucidate entirely new mechanisms for evolution of CHD risk and pathogenesis of CHD.

Insights from rodent models

The use of rodent models to dissect the complex genetic etiology of atherosclerosis offers a powerful alternative to human studies because experimental parameters, such as environment, breeding scheme, and detailed phenotyping, can be more easily controlled. Over the last two decades, linkage and gene targeting studies in rodents have provided profound genetic insights into the pathogenesis of dyslipidemia and atherosclerosis (Allayee *et al.*, 2003; Glass *et al.*, 2001; Lusis, 2000; Mehrabian *et al.*, 2002). Indeed, some of these studies have led to the identification of genes that would not have otherwise been considered as candidates for atherosclerosis. However, rodents rarely develop atheroma of the coronary arteries and very few models go on to develop myocardial infarction as a result of atherosclerotic plaque rupture – the single most important pathophysiological endpoint associated with CHD (Rekhter, 2002). Therefore, potential translational implications derived from rodent studies should always be confirmed in human subjects.

Gene targeting studies

Before the advent of the techniques of genetic manipulation it was well known that different inbred strains had varying degrees of susceptibility to diet-induced atherosclerosis, but the arterial lesions induced by even extreme diets were limited in size, complexity, and distribution. This changed when efforts to identify genetic modifiers of dyslipidemia and atherosclerosis in mice were performed through gene targeting using loss-of-function or gain-of-function approaches with the production of several gene knockout and transgenic mice.

Early loss-of-function studies focused on molecules previously identified as potentially important modifiers of lipid handling. From these studies, *apoE* (Plump *et al.*, 1992; Zhang *et al.*, 1992) and the LDL receptor (Ishibashi *et al.*, 1993) were confirmed as important modifiers of cholesterol levels. Mice lacking either *apoE* or the LDL receptor develop significant atherosclerotic lesions when fed a high fat diet. This observation paved the way for a deluge of studies looking at genetic influences on atherosclerosis lesion size in *apoE* and LDL receptor knockout mice (Glass and Witztum, 2001). Consequently, a solid body of new information has emerged on the mechanisms regulating plasma lipoprotein levels and controlling the initial stages of atherogenesis. The obvious advantage of studying the effect of a gene in mice predisposed to developing atherosclerosis is that anti-atherogenic effects of a gene may be more readily detected (Table 24.3).

Gain-of-function experiments have been used extensively to determine the biological effect of genes on lipid profile and atherosclerosis. However, gain-of-function approaches are a less elegant genetic tool for determining a gene's effect. This is because in transgenic animals the gene expressed may be derived from a different species, effects due to insertion of the transgene are not accounted for, gene expression may be unphysiologically high and developmental effects may be important. Despite these drawbacks, many insightful observations as to the genetic modifiers of lipid metabolism and atherosclerosis have been identified (Table 24.3).

Linkage studies

QTL studies of atherosclerosis with experimental crosses of a wide variety of inbred strains of mice, maintained on either atherogenic diets or bred onto sensitized backgrounds (see below) have identified a large number of loci with highly significant evidence of linkage, as reviewed by Allayee and colleagues (Allayee *et al.*, 2003). Interestingly, many, but not all, (Bodnar *et al.*, 2002) of the novel loci identified in mice do not associate with determinants of plasma lipid levels. An example of this is a gene that was localized to chromosome 6 in mice at the Artles locus, a locus associated with protection against atherosclerosis (Mehrabian *et al.*, 2001). The gene, derived from CAST/Ei mice was subsequently identified as a loss of function mutation of 5-lipoxygenase (5-LO) (Kuhn *et al.*, 2003; Mehrabian *et al.*, 2002). This enzyme regulates leukotriene production and has significant effects on atherosclerotic lesions, independent of cholesterol levels, highlighting the current emphasis on non-lipid, inflammatory modifiers of CHD.

Studies in rats have identified a number of chromosomal loci associated with cardiovascular risk. The spontaneously hypertensive rat (SHR) is insulin resistant and a model of the human insulin-resistance syndromes. In one study, QTLs for SHR defects in glucose and fatty acid metabolism, hypertriglyceridemia and hypertension mapped to a single locus on rat chromosome 4. Following this the combined use of cDNA microarrays, congenic mapping and radiation hybrid mapping identified a defective SHR gene, Cd36 (also known as Fat, as it encodes fatty acid translocase), at the peak of linkage to these QTLs (Aitman *et al.*, 1999). Subsequent sequencing of CD36 in the SHR showed significant sequence variation implying functional deficiency of CD36. However, deletion of CD36 in mice paradoxically protects from atherosclerosis (Febbraio *et al.*, 2000) and is likely

Table 24.3. Examples of loss-of-function or gain-of-function studies on the background of inbred mouse strains or atherosclerosis susceptibility strains

Gene targeting	Gene	Effect of atherosclerosis	Biological effect	Reference
Direct gene targeting				
Loss-of-function	*ApoE*	+++[a]	↑ LDL	(Plump *et al.*, 1992)
	LDL R	+++	↑ LDL	(Ishibashi *et al.*, 1993)
	eNOS	−	↓ LDL oxidation	(Shi *et al.*, 2002)
Gain-of-function	*CETP*	++	↑ LDL	(Marotti *et al.*, 1993)
	ApoAI	−	↑ HDL	(Walsh *et al.*, 1989)
	ApoAII	++	Larger HDL	(Warden *et al.*, 1993)
	ApoB100	++	↑ LDL	(Linton *et al.*, 1992)
	ApoCI	++	↑ cholesterol	(Shachter *et al.*, 1996)
	ApoCII	+/−	↑ VLDL/TG	(Shachter *et al.*, 1994)
	ApoCIII	+/−	↑ TG	(Ito *et al.*, 1990)
	LCAT	++	↑ HDL	(Berard *et al.*, 1997)
Gene targeting to susceptible backgrounds				
Loss-of-function	*15-LO*	−	↓ Oxidation	(Cyrus *et al.*, 1999)
	CD154	−	↓ CD40 signaling	(Lutgens *et al.*, 1999)
	5-LO	−	↓ inflammation	(Mehrabian *et al.*, 2002)
	CD36	−	↓ cholesterol uptake, ↑ cholesterol efflux and ↑ TG	(Febbraio *et al.*, 2000)
	CXCR2	−	↓ macrophage infiltrate	(Boisvert *et al.*, 1997)
	CX3CR1	−	↓ macrophage infiltrate	(Lesnik *et al.*, 2003)
	CCR2	−	↓ macrophage infiltrate	(Boring *et al.*, 1998)
	MCP1	−	↓ macrophage infiltrate	(Gu *et al.*, 1998)
	PON	+	↓ removal oxidised LDL	(Shih *et al.*, 2000)
	ApoA1	+	↓ HDL cholesterol	(Voyiaziakis *et al.*, 1998)
	eNOS	++	↓ NO production	(Kuhlencordt *et al.*, 2001)
	M-CSF	−	↓ macrophage infiltrate	(Smith *et al.*, 1995)
Gain-of-function	*15-LO*	++	↑ LDL oxidation	(Harats *et al.*, 2000)
	ApoAI	−	Increased HDL	(Benoit *et al.*, 1999)
	MMP-1	−	↓ remodeling	(Lemaitre *et al.*, 2001)

[a] The number of + or − denotes relative pro- or anti-atherogenic effect

Abbreviations: apolipoprotein (Apo), low density lipoprotein (LDL), endothelial (e) nitric oxide synthase (NOS), cholesteryl ester transfer protein (CETP), lecithin:cholesterol acyltransferase (LCAT), lipoxygenase (LO), Monocyte chemoattractant protein-1 (MCP-1), Paraoxonase (PON), inducible NOS (iNOS), macrophage colony-stimulating factor (M-CSF), matrix metalloproteinase (MMP) (Adapted from Glass and Wiztum, 2001).

to be due to loss of CD36 from macrophages, which reduces their capacity to take up cholesterol. This observation delineates an important principle relevant to other genetic studies in that the genetic mutation (loss of CD36), causally related to development of an atherogenic metabolic milieu, paradoxically protects against the development of atherosclerosis. This kind of ''functional genomic'' approach may be the only route to a full understanding of the mechanistic and

pathophysiological role(s) of predisposing susceptibility genes.

Conclusions

The current data indicate that a large number of individuals are at an increased risk of developing CHD because they carry several predisposing genetic variants, leading to risk accumulation. In common with other complex diseases, identification of the susceptibility genes that account for this risk has proved difficult. Many investigators have chosen to study candidate genes but, with the exception of certain polymorphisms in genes regulating lipid metabolism, the role of most gene polymorphisms is controversial or unknown. Genome–wide linkage studies have robustly mapped a small number of loci for CHD but most have as yet not led to identification of the underlying disease gene. Given the rapid advances taking place in very high throughput genotyping, genome sequence annotation and haplotype mapping, substantial new progress in this area can be expected in the next two to three years.

A major part of our understanding of atherosclerosis pathogenesis has arisen from the genetic manipulation of mice as a means to understanding molecular and cellular mechanisms. Although rodent models only develop early atherosclerotic lesions, even after genetic and dietary manipulation, these models have provided a means to study gene–gene and gene–environment interactions. Importantly, rodent models of CHD may also be used to specifically evaluate therapeutics and preventative interventions. In addition, studies of naturally occurring genetic variations in rodents predisposing to atherosclerosis have great potential to reveal new genes and pathways important in familial and population susceptibility to CHD. To date, many of the loci for atherosclerosis-related traits in rodents have been mapped, but few genes identified. With important progress being made in rodent genomics and bioinformatics the process of gene identification should accelerate.

REFERENCES

Abifadel, M., Varret, M., Rabes, J.P. *et al.* (2003). Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat Genet*, **34**, 154–6.

Aitman, T.J., Dong, R., Vyse, T.J. *et al.* (2006). Copy number polymorphism in Fcgr3 predisposes to glomerulonephritis in rats and humans. *Nature*, **439**, 798–9.

Aitman, T.J., Glazier, A.M., Wallace, C.A. *et al.* (1999). Identification of Cd36 (Fat) as an insulin-resistance gene causing defective fatty acid and glucose metabolism in hypertensive rats. *Nat Genet*, **21**, 76–83.

Aitman, T.J., Godsland, I.F., Farren, B. *et al.* (1997*a*). Defects of insulin action on fatty acid and carbohydrate metabolism in familial combined hyperlipidemia. *Arterioscler Thromb Vasc Biol*, **17**, 748–54.

Aitman, T.J., Gotoda, T., Evans, A.L. *et al.* (1997*b*). Quantitative trait loci for cellular defects in glucose and fatty acid metabolism in hypertensive rats. *Nat Genet*, **16**, 197–201.

Allayee, H., Ghazalpour, A. and Lusis, A.J. (2003). Using mice to dissect genetic factors in atherosclerosis. *Arterioscler Thromb Vasc Biol*, **23**, 1501–9.

Altshuler, D., Hirschhorn, J.N., Klannemark, M. *et al.* (2000). The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat Genet*, **26**, 76–80.

Altshuler, D. and Hirschhorn, J.N. (2005). MEF2A sequence variants and coronary artery disease: a change of heart? *J Clin Invest*, **4**, 831–3.

Aouizerat, B.E., Allayee, H., Cantor, R.M. *et al.* (1999). A genome scan for familial combined hyperlipidemia reveals evidence of linkage with a locus on chromosome 11. *Am J Hum Genet*, **65**, 397–412.

Arca, M., Zuliani, G., Wilund, K. *et al.* (2002). Autosomal recessive hypercholesterolaemia in Sardinia, Italy, and mutations in ARH: a clinical and molecular genetic analysis. *Lancet*, **359**, 841–7.

Austin, M.A. (1992). Genetic epidemiology of low-density lipoprotein subclass phenotypes. *Ann Med*, **24**, 477–81.

Babirak, S.P., Brown, B.G. and Brunzell, J.D. (1992). Familial combined hyperlipidemia and abnormal lipoprotein lipase. *Arterioscler Thromb*, **12**, 1176–83.

Benoit, P., Emmanuel, F., Caillaud, J.M. *et al.* (1999). Somatic gene transfer of human ApoA-I inhibits atherosclerosis progression in mouse models. *Circulation*, **99**, 105–10.

Berard, A. M., Foger, B., Remaley, A. *et al.* (1997). High plasma HDL concentrations associated with enhanced atherosclerosis in transgenic mice overexpressing lecithin-cholesteryl acyltransferase. *Nat Med*, **3**, 744–9.

Berge, K. E., Tian, H., Graf, G. A. *et al.* (2000). Accumulation of dietary cholesterol in sitosterolemia caused by mutations in adjacent ABC transporters. *Science*, **290**, 1771–5.

BHF Family Heart Study Research Group. (2005). A Genomewide Linkage Study of 1,933 Families Affected by Premature Coronary Artery Disease: The British Heart Foundation (BHF) Family Heart Study. *Am J Hum. Genet.*, **77**, 1011–20.

Bodnar, J. S., Chatterjee, A., Castellani, L. W. *et al.* (2002). Positional cloning of the combined hyperlipidemia gene Hyplip1. *Nat Genet*, **30**, 110–6.

Boisvert, W. A., Spangenberg, J. and Curtiss, L. K. (1997). Role of leukocyte-specific LDL receptors on plasma lipoprotein cholesterol and atherosclerosis in mice. *Arterioscler Thromb Vasc Biol*, **17**, 340–7.

Boring, L., Gosling, J., Cleary, M. and Charo, I. F. (1998). Decreased lesion formation in CCR2-/- mice reveals a role for chemokines in the initiation of atherosclerosis. *Nature*, **394**, 894–7.

Broeckel, U., Hengstenberg, C., Mayer, B. *et al.* (2002). A comprehensive linkage analysis for myocardial infarction and its related risk factors. *Nat Genet*, **30**, 210–4.

Cabezas, M. C., de Bruin, T. W., Jansen, H. *et al.* (1993). Impaired chylomicron remnant clearance in familial combined hyperlipidemia. *Arterioscler Thromb*, **13**, 804–14.

Cambien, F., Poirier, O., Lecerf, L. *et al.* (1992). Deletion polymorphism in the gene for angiotensin-converting enzyme is a potent risk factor for myocardial infarction. *Nature*, **359**, 641–4.

Chiodini, B. D. and Lewis, C. M. (2003). Meta-analysis of 4 coronary heart disease genome-wide linkage studies confirms a susceptibility locus on chromosome 3q. *Arterioscler Thromb Vasc Biol*, **23**, 1863–8.

Cullen, P., Farren, B., Scott, J. and Farrall, M. (1994). Complex segregation analysis provides evidence for a major gene acting on serum triglyceride levels in 55 British families with familial combined hyperlipidemia. *Arterioscler Thromb*, **14**, 1233–49.

Cyrus, T., Witztum, J. L., Rader, D. J. *et al.* (1999). Disruption of the 12/15-lipoxygenase gene diminishes atherosclerosis in apo E-deficient mice. *J Clin Invest*, **103**, 1597–604.

Dallinga-Thie, G. M., Bu, X. D., Linde-Sibenius, T. M. *et al.* (1996). Apolipoprotein *A-I/C-III/A-IV* gene cluster in familial combined hyperlipidemia: effects on LDL-cholesterol and apolipoproteins B and C-III. *J Lipid Res*, **37**, 136–47.

Doggen, C. J., Manger, C. V., Bertina, R. M. *et al.* (1998). A genetic propensity to high factor VII is not associated with the risk of myocardial infarction in men. *Thromb Haemost*, **80**, 281–5.

Durrington, P. (2003). Dyslipidaemia. *Lancet*, **362**,717–31.

Durrington, P. N. (2003). Lipid and lipoprotein disorders. In Warrell, D. A., Cox, T. M., Firth, J. D. and Benz, E. J., eds., *Oxford textbook of medicine*, 4th edn. Oxford: Oxford University Press, pp. 74–90.

Eden, E. R., Patel, D. D., Sun, X. M. *et al.* (2002). Restoration of LDL receptor function in cells from patients with autosomal recessive hypercholesterolemia by retroviral expression of *ARH1*. *J Clin Invest*, **110**, 1695–702.

Eichenbaum-Voline, S., Olivier, M., Jones, E. L. *et al.* (2004). Linkage and association between distinct variants of the *APOA1/C3/A4/A5* gene cluster and familial combined hyperlipidemia. *Arterioscler Thromb Vasc Biol*, **24**, 167–74.

Febbraio, M., Podrez, E. A., Smith, J. D. *et al.* (2000). Targeted disruption of the class B scavenger receptor CD36 protects against atherosclerotic lesion development in mice. *J Clin Invest*, **105**, 1049–56.

Francke, S., Manraj, M., Lacquemant, C. *et al.* (2001). A genome-wide scan for coronary heart disease suggests in Indo-Mauritians a susceptibility locus on chromosome 16p13 and replicates linkage with the metabolic syndrome on 3q27. *Hum Mol Genet*, **10**, 2751–65.

Furuhashi, M., Ura, N., Nakata, T. and Shimamoto, K. (2003). Insulin sensitivity and lipid metabolism in human CD36 deficiency. *Diabetes Care*, **26**, 471–4.

Fuster, V. (1999). Epidemic of cardiovascular disease and stroke: the three main challenges. Presented at the 71st Scientific Sessions of the American Heart Association. Dallas, Texas. *Circulation*, **99**, 1132–7.

Garcia, C. K., Wilund, K., Arca, M. *et al.* (2001). Autosomal recessive hypercholesterolemia caused by mutations in a putative LDL receptor adaptor protein. *Science*, **292**, 1394–8.

Girelli, D., Russo, C., Ferraresi, P. *et al.* (2000). Polymorphisms in the factor VII gene and the risk of myocardial infarction in patients with coronary artery disease. *N Engl J Med*, **343**, 774–80.

Glass, C. K. and Witztum, J. L. (2001). Atherosclerosis. the road ahead. *Cell*, **104**, 503–16.

Goldbourt, U., Yaari, S. and Medalie, J. H. (1997). Isolated low HDL cholesterol as a risk factor for coronary heart disease mortality. A 21-year follow-up of 8000 men. *Arterioscler Thromb Vasc Biol*, **17**, 107–13.

Goldstein, J., Hazzard, W. R., Schrott, H. G., Bierman, E. L. and Motulsky, A. G. (1973). Hyperlipidemia in coronary heart disease. I. Lipid levels in 500 survivors of myocardial infarction. *J Clin Invest*, **52**, 1533–43.

Goldstein, J., Hobbs, H. and Brown, M. (2001). Familial hypercholesterolemia. In Scriver, C. R., Beaudet, A. I., Sly, W. S. *et al.* (eds.), *The metabolic and molecular bases of inherited disease* 8th edn. McGraw Hill, pp. 2863–913.

Gonzalez, E. *et al.* (2005). The influence of *CCL3L1* gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science*, **307**, 1434–40.

Gu, L., Okada, Y., Clinton, S. K. *et al.* (1998). Absence of monocyte chemoattractant protein-1 reduces atherosclerosis in low density lipoprotein receptor-deficient mice. *Mol Cell*, **2**, 275–81.

Harats, D., Shaish, A., George, J. *et al.* (2000). Overexpression of 15-lipoxygenase in vascular endothelium accelerates early atherosclerosis in LDL receptor-deficient mice. *Arterioscler Thromb Vasc Biol*, **20**, 2100–5.

Harrap, S. B., Zammit, K. S., Wong, Z. Y. *et al.* (2002). Genome-wide linkage analysis of the acute coronary syndrome suggests a locus on chromosome 2. *Arterioscler Thromb Vasc Biol*, **22**, 874–8.

Hawe, E., Talmud, P. J., Miller, G. J. and Humphries, S. E. (2003). Family history is a coronary heart disease risk factor in the Second Northwick Park Heart Study. *Ann Hum Genet*, **67**, 97–106.

Helgadottir, A., Manolescu, A., Thorleifsson, G. *et al.* (2004). The gene encoding 5-lipooxygenase activating protein (FLAP) confers risk of myocardial infarction and stroke. *Nature Genetics*, **36**, 233–9.

Helgadottir, A., Manolescu, A., Helgason, A. *et al.* (2006). A variant of the gene encoding leukotriene A4 hydrolase confers ethnicity-specific risk of myocardial infarction. *Nature Genetics*, **38**, 68–74.

Hill, J. S., Hayden, M. R., Frohlich, J. and Pritchard, P. H. (1991). Genetic and environmental factors affecting the incidence of coronary artery disease in heterozygous familial hypercholesterolemia. *Arterioscler Thromb*, **11**, 290–7.

Iacoviello, L., Di Castelnuovo, A., De Knijff, P. *et al.* (1998). Polymorphisms in the coagulation factor VII gene and the risk of myocardial infarction. *N Engl J Med*, **338**, 79–85.

Ishibashi, S., Brown, M. S., Goldstein, J. L. *et al.* (1993). Hypercholesterolemia in low density lipoprotein receptor knockout mice and its reversal by adenovirus-mediated gene delivery. *J Clin Invest*, **92**, 883–93.

Ito, Y., Azrolan, N., O'Connell, A., Walsh, A. and Breslow, J. L. (1990). Hypertriglyceridemia as a result of human apo CIII gene expression in transgenic mice. *Science*, **249**, 790–3.

Kissebah, A. H., Alfarsi, S. and Evans, D. J. (1984). Low density lipoprotein metabolism in familial combined hyperlipidemia. Mechanism of the multiple lipoprotein phenotypic expression. *Arteriosclerosis*, **4**, 614–24.

Kuhlencordt, P. J., Gyurko, R., Han, F. *et al.* (2001). Accelerated atherosclerosis, aortic aneurysm formation, and ischemic heart disease in apolipoprotein E/endothelial nitric oxide synthase double-knockout mice. *Circulation*, **104**, 448–54.

Kuhn, H., Anton, M., Gerth, C. and Habenicht, A. (2003). Amino acid differences in the deduced 5-lipoxygenase sequence of CAST atherosclerosis-resistance mice confer impaired activity when introduced into the human ortholog. *Arterioscler Thromb Vasc Biol*, **23**, 1072–6.

Lane, A., Green, F., Scarabin, P. Y. *et al.* (1996). Factor VII Arg/Gln353 polymorphism determines factor VII coagulant activity in patients with myocardial infarction (MI) and control subjects in Belfast and in France but is not a strong indicator of MI risk in the ECTIM study. *Atherosclerosis*, **119**, 119–27.

Lee, M. H., Lu, K., Hazard, S. *et al.* (2001). Identification of a gene, ABCG5, important in the regulation of dietary cholesterol absorption. *Nat Genet*, **27**, 79–83.

Lemaitre, V., O'Byrne, T. K., Borczuk, A. C. *et al.* (2001). *ApoE* knockout mice expressing human matrix metalloproteinase-1 in macrophages have less advanced atherosclerosis. *J Clin Invest*, **107**, 1227–34.

Lesnik, P., Haskell, C. A. and Charo, I. F. (2003). Decreased atherosclerosis in *CX3CR1*-/- mice reveals a role for fractalkine in atherogenesis. *J Clin Invest*, **111**, 333–40.

Lindpaintner, K., Pfeffer, M. A., Kreutz, R. *et al.* (1995). A prospective evaluation of an angiotensin-converting-enzyme gene polymorphism and the risk of ischemic heart disease. *N Engl J Med*, **332**, 706–11.

Linton, M. F., Pierotti, V. and Young, S. G. (1992). Reading-frame restoration with an apolipoprotein B gene frameshift mutation. *Proc Natl Acad Sci USA*, **89**, 11431–5.

Lusis, A. J. (2000). Atherosclerosis. *Nature*, **407**, 233–41.

Lusis, A. J. (2003). Genetic factors in cardiovascular disease. 10 questions. *Trends Cardiovasc Med*, **13**, 309–16.

Lutgens, E., Gorelik, L., Daemen, M. J. *et al.* (1999). Requirement for CD154 in the progression of atherosclerosis. *Nat Med*, **5**, 1313–16.

Mahley, R. W. and Rall, S. C. Jr. (2001). Type III hyperlipoproteinemia (dysbetalipoproteinemia): the role of apolipoprotein E in normal and abnormal lipoprotein metabolism. In Scriver, C. R., Beaudet, A. I., Sly, W. S. *et al.* (eds.), *The metabolic and molecular basis of inherited diseases*, 8th edn. New York: McGraw Hill, pp. 2835–62.

Marcil, M., Brooks-Wilson, A., Clee, S. M. *et al.* (1999). Mutations in the *ABC1* gene in familial HDL deficiency with defective cholesterol efflux. *Lancet*, **354**, 1341–6.

Marks, D., Thorogood, M., Neil, H. A. and Humphries, S. E. (2003). A review on the diagnosis, natural history, and treatment of familial hypercholesterolaemia. *Atherosclerosis*, **168**, 1–14.

Marotti, K. R., Castle, C. K., Boyle, T. P. *et al.* (1993). Severe atherosclerosis in transgenic mice expressing simian cholesteryl ester transfer protein. *Nature*, **364**, 73–5.

McKeigue, P. M., Ferrie, J. E., Pierpoint, T. and Marmot, M. G. (1993). Association of early-onset coronary heart disease in South Asian men with glucose intolerance and hyperinsulinemia. *Circulation*, **87**, 152–61.

Mehrabian, M., Wong, J., Wang, X. *et al.* (2001). Genetic locus in mice that blocks development of atherosclerosis despite extreme hyperlipidemia. *Circ Res*, **89**, 125–30.

Mehrabian, M., Allayee, H., Wong, J. *et al.* (2002). Identification of 5-lipoxygenase as a major gene contributing to atherosclerosis susceptibility in mice. *Circ Res*, **91**, 120–6.

Miyaoka, K., Kuwasako, T., Hirano, K. *et al.* (2001). CD36 deficiency associated with insulin resistance. *Lancet*, **357**, 686–7.

Motulsky, A. G. and Brunzell, J. D. (2002). Genetics of Coronary Atherosclerosis. In King, R. A., Rotter, J. I. and Motulsky, A. G. (eds.), *The genetic basis of common diseases*, 2nd edn. Oxford: Oxford University Press, pp. 105–26.

Murray, C. J. and Lopez, A. D. (1997). Alternative projections of mortality and disability by cause 1990–2020: Global Burden of Disease Study. *Lancet*, **349**, 1498–504.

Myant, N. B. (1993). Familial defective apolipoprotein B-100: a review, including some comparisons with familial hypercholesterolaemia. *Atherosclerosis*, **104**, 1–18.

Nagano, M., Yamashita, S., Hirano, K. *et al.* (2002). Two novel missense mutations in the CETP gene in Japanese hyperalphalipoproteinemic subjects: high-throughput assay by Invader assay. *J Lipid Res*, **43**, 1011–8.

Naoumova, R. P., Bonney, S. A., Eichenbaum-Voline, S. *et al.* (2003). Confirmed Locus on Chromosome 11p and Candidate Loci on 6q and 8p for the triglyceride and cholesterol traits of combined hyperlipidemia. *Arterioscler Thromb Vasc Biol*, **23**, 2070–7.

Naoumova, R. P., Neuwirth, C., Lee, P. *et al.* (2004). Autosomal recessive hypercholesterolaemia: long-term follow up and response to treatment. *Atherosclerosis*, **174**, 165–72.

Naoumova, R. P., Tosi, I., Patel, D. *et al.* (2005). Severe hypercholesterolemia in four British families with the D374Y mutation in the *PCSK9* gene: long-term follow-up and treatment response. *Arterioscler Thromb Vasc Biol*, **25**, 2654–60.

Navarro-Lopez, F. (2002). Genes and coronary heart disease. *Rev Esp Cardiol*, **55**, 413–31.

Oram, J. F. (2002). ATP-binding cassette transporter A1 and cholesterol trafficking. *Curr Opin Lipidol*, **13**, 373–81.

Ozaki, K., Inoue, K., Sato, H. *et al.* (2004). Functional variation in LGALS2 confers risk of myocardial infarction and regulates lymphotoxin-alpha secretion in vitro. *Nature*, **429**, 72–5.

Ozaki, K., Ohnishi, Y., Iida, A. *et al.* (2002). Functional SNPs in the lymphotoxin-alpha gene that are associated with susceptibility to myocardial infarction. *Nature Genetics*, **32**, 650–4.

Packard, C. J. and Shepherd, J. (1999). Physiology of the lipoprotein transport system: an overview of lipoprotein metabolism. In Betteridge, D. J., Illingworth, D. R. and Shepherd, J. (eds.), *Lipoproteins in health and disease*, 1st edn. London: Arnold.

Pajukanta, P., Terwilliger, J. D., Perola, M. *et al.* (1999). Genomewide scan for familial combined hyperlipidemia genes in finnish families, suggesting multiple susceptibility loci influencing triglyceride, cholesterol, and apolipoprotein B levels. *Am J Hum Genet*, **64**, 1453–63.

Pajukanta, P., Cargill, M., Viitanen, L. *et al.* (2000). Two loci on chromosomes 2 and X for premature coronary heart disease identified in early- and late-settlement populations of Finland. *Am J Hum Genet*, **67**, 1481–93.

Pennacchio, L.A., Olivier, M., Hubacek, J.A. *et al.* (2001). An apolipoprotein influencing triglycerides in humans and mice revealed by comparative sequencing. *Science*, **294**, 169–73.

Pimstone, S.N., Sun, X.M., du, S.C. *et al.* (1998). Phenotypic variation in heterozygous familial hypercholesterolemia: a comparison of Chinese patients with the same or similar mutations in the LDL receptor gene in China or Canada. *Arterioscler Thromb Vasc Biol*, **18**, 309–15.

Plump, A.S., Smith, J.D., Hayek, T. *et al.* (1992). Severe hypercholesterolemia and atherosclerosis in apolipoprotein E-deficient mice created by homologous recombination in ES cells. *Cell*, **71**, 343–53.

Pravenec, M., Landa, V., Zidek, V. *et al.* (2001). Transgenic rescue of defective Cd36 ameliorates insulin resistance in spontaneously hypertensive rats. *Nature Genetics*, **27**, 156–8.

Pullinger, C.R., Eng, C., Salen, G. *et al.* (2002). Human cholesterol 7α-hydroxylase (CYP7A1) deficiency has a hypercholesterolemic phenotype. *J Clin Invest*, **110**, 109–17.

Reaven, G.M. (1988). Banting lecture 1988. Role of insulin resistance in human disease. *Diabetes*, **37**, 1595–607.

Rekhter, M.D. (2002). How to evaluate plaque vulnerability in animal models of atherosclerosis? *Cardiovasc Res*, **54**, 36–41.

Salen, G., Shefer, S., Nguyen, L. *et al.* (1992). Sitosterolemia. *J Lipid Res*, **33**, 945–55.

Samani, N.J., Boultby, R., Butler, R., Thompson, J.R. and Gooddall, A.H. (2001). Telomere shortening in atherosclerosis. *Lancet*, **358**, 472–3.

Schaefer, J.R., Scharnagl, H., Baumstark, M.W. *et al.* (1997). Homozygous familial defective apolipoprotein B-100. Enhanced removal of apolipoprotein E-containing VLDLs and decreased production of LDLs. *Arterioscler Thromb Vasc Biol*, **17**, 348–53.

Seidah, N.G., Benjannet, S., Wickham, L. *et al.* (2003). The secretory proprotein convertase neural apoptosis-regulated convertase 1 (NARC-1): liver regeneration and neuronal differentiation. *Proc Natl Acad Sci USA*, **100**, 928–33.

Shachter, N.S., Hayek, T., Leff, T. *et al.* (1994). Overexpression of apolipoprotein CII causes hypertriglyceridemia in transgenic mice. *J Clin Invest*, **93**, 1683–90.

Shachter, N.S., Ebara, T., Ramakrishnan, R. *et al.* (1996). Combined hyperlipidemia in transgenic mice overexpressing human apolipoprotein Cl. *J Clin Invest*, **98**, 846–55.

Shi, W., Wang, X., Shih, D.M. *et al.* (2002). Paradoxical reduction of fatty streak formation in mice lacking endothelial nitric oxide synthase. *Circulation*, **105**, 2078–82.

Shih, D.M., Xia, Y.R., Wang, X.P. *et al.* (2000). Combined serum paraoxonase knockout/apolipoprotein E knockout mice exhibit increased lipoprotein oxidation and atherosclerosis. *J Biol Chem*, **275**, 17527–35.

Shoulders, C.C., Jones, E.L. and Naoumova, R.P. (2004). Genetics of familial combined hyperlipidemia and risk of coronary heart disease. *Hum Mol Genet*, **13** Spec No 1, R149–R160.

Slack, J. (1969). Risks of ischaemic heart-disease in familial hyperlipoproteinaemic states. *Lancet*, **2**, 1380–2.

Smith, J.D., Trogan, E., Ginsberg, M. *et al.* (1995). Decreased atherosclerosis in mice deficient in both macrophage colony-stimulating factor (op) and apolipoprotein E. *Proc Natl Acad Sci USA*, **92**, 8264–8.

Sorensen, T.I., Nielsen, G.G., Andersen, P.K. and Teasdale, T.W. (1988). Genetic and environmental influences on premature death in adult adoptees. *N Engl J Med*, **318**, 727–32.

Soria, L.F., Ludwig, E.H., Clarke, H.R., Vega, G.L., Grundy, S.M. and McCarthy, B.J. (1989). Association between a specific apolipoprotein B mutation and familial defective apolipoprotein B-100. *Proc Natl Acad Sci USA*, **86**, 587–91.

Sun, X.M., Eden, E.R., Tosi, I. *et al.* (2005). Evidence for effect of mutant PCSK9 on apolipoprotein B secretion as the cause of unusually severe dominant hypercholesterolaemia. *Hum Mol Genet*, **14**, 1161–9.

Tall, A.R. (1998). An overview of reverse cholesterol transport. *Eur Heart J*, **19** Suppl A, A31–A35.

Thompson, G.R. (1999). Familial hypercholesterolaemia. In Betteridge, D.J., Illingworth, D.R. and Shepherd, J. (eds.), *Lipoproteins in health and disease*, 1st edn. London: Arnold.

Tontonoz, P., Nagy, L., Alvarez, J.G., Thomazy, V.A. and Evans, R.M. (1998). PPARgamma promotes monocyte/macrophage differentiation and uptake of oxidized LDL. *Cell*, **93**, 241–52.

Topol, E.J. (2005). The genomic basis of myocardial infarction. *Journal of the American College of Cardiology*, **46**, 1456–65.

Varret, M., Rabes, J. P., Saint-Jore, B. *et al.* (1999). A third major locus for autosomal dominant hypercholesterolemia maps to 1p34.1-p32. *Am J Hum Genet*, **64**, 1378–87.

Venkatesan, S., Cullen, P., Pacy, P., Halliday, D. and Scott, J. (1993). Stable isotopes show a direct relation between VLDL apoB overproduction and serum triglyceride levels and indicate a metabolically and biochemically coherent basis for familial combined hyperlipidemia. *Arterioscler Thromb*, **13**, 1110–18.

von Eckardstein, A., Nofer, J. R. and Assmann, G. (2001). High density lipoproteins and arteriosclerosis. Role of cholesterol efflux and reverse cholesterol transport. *Arterioscler Thromb Vasc Biol*, **21**, 13–27.

Voyiaziakis, E., Goldberg, I. J., Plump, A. S. *et al.* (1998). ApoA-I deficiency causes both hypertriglyceridemia and increased atherosclerosis in human apoB transgenic mice. *J Lipid Res*, **39**, 313–21.

Walsh, A., Ito, Y. and Breslow, J. L. (1989). High levels of human apolipoprotein A-I in transgenic mice result in increased plasma levels of small high density lipoprotein (HDL) particles comparable to human HDL3. *J Biol Chem*, **264**, 6488–94.

Wang, L., Fan, C., Topol, S. E., Topol, E. J. and Wang, Q. (2003). Mutation of *MEF2A* in an inherited disorder with features of coronary artery disease. *Science*, **302**, 1578–81.

Warden, C. H., Hedrick, C. C., Qiao, J. H., Castellani, L. W. and Lusis, A. J. (1993). Atherosclerosis in transgenic mice overexpressing apolipoprotein A-II. *Science*, **261**, 469–72.

Zdravkovic, S., Wienke, A., Pedersen, N. L. *et al.* (2002). Heritability of death from coronary heart disease: a 36-year follow-up of 20 966 Swedish twins. *J Intern Med*, **252**, 247–54.

Zhang, S. H., Reddick, R. L., Piedrahita, J. A. and Maeda, N. (1992). Spontaneous hypercholesterolemia and arterial lesions in mice lacking apolipoprotein E. *Science*, **258**, 468–71.

25

Genetics of hypertension

B. Keavney and M. Lathrop

Introduction

Higher levels of blood pressure are a major risk factor for coronary heart disease (CHD), stroke, cardiac failure and renal failure. Meta-analysis of individual participant data from over 1 million people in prospective cohort studies has shown that the log risk of coronary heart disease and stroke are linearly related to the level of blood pressure throughout a range which extends well into that usually regarded as ''normal,'' with no evidence of a threshold below which blood pressure no longer influences risk (Lewington *et al.*, 2002). Studies involving participants from many different countries show that high blood pressure is a risk for these conditions throughout the world, with the expected impact of high blood pressure set to increase as developing countries industrialise (Yusuf *et al.*, 2004). Blood pressure is distributed approximately log-normally in all populations studied to date; within this distribution, cut-off criteria for levels of blood pressure which confer a level of risk requiring treatment (and thus define conventional hypertension) have changed over time, in general being revised downward as the importance of even relatively small elevations in blood pressure has been more widely appreciated. Currently, the European Hypertension Society guidelines define Grade I hypertension as a systolic pressure of 140–159 mmHg and/or a diastolic pressure of 90–99 mmHg; such a definition would result in some 40% of males and 33% of females in the UK carrying a diagnosis of hypertension (European Society of Hypertension/European Society of Cardiology Guidelines Committee, 2003). It is now recognised that multiple drugs are usually required for adequate control of hypertension, but the practicalities of devising and delivering a multidrug regime that is effective and well tolerated by individual patients remains a significant challenge; it has been estimated that the resulting undertreatment of hypertension costs 125 600 CHD and stroke events (half of which are fatal) each year in the UK (He and MacGregor, 2003). Thus, additional drug targets for hypertension identified through genetics could be of major public health importance, as could common genetic variants having only small influences on blood pressure levels within the generally accepted ''normal range.'' A number of specific pathologies, mostly involving the kidney (such as diabetic nephropathy, glomerulonephritis, and renal artery stenosis) are causes of ''secondary'' hypertension. While these are important, and may have significant genetic inputs in their own right, the remainder of our discussion relates to ''primary,'' also called ''essential'' hypertension, in which no specific predisposing cause has been found (representing over 90% of all prevalent cases).

Blood pressure exhibits a relatively modest familial correlation, a portion of which can be attributed to genetic factors. Epidemiological studies have led to estimates of 20–40% for the heritability of blood pressure, and quantitative genetic analysis supports the hypothesis that blood pressure is modulated by action of a large

number of genes (Rudan *et al.*, 2003). The feasibility of mapping and identifying any of these genes is possible only if a small number contribute disproportionately; this remains unproven but it is not an unreasonable hypothesis based on present knowledge of quantitative trait genetics (Farrall, 2004). Single measurements of blood pressure are subject to substantial error; some investigators have attempted to minimize this by the use of ambulatory blood pressure monitors, which record blood pressure at multiple time points throughout the day. In general such studies have provided higher estimates of heritability, suggesting that this approach may be of considerable value in maximising the genetic ''signal'' (Fagard *et al.*, 1995; Fava *et al.*, 2004; Kupper *et al.*, 2005). However, while several Mendelian forms of hypertension (all with low individual population prevalence, together contributing <1% to the population burden of disease) have been successfully mapped, as yet no single gene with a large contribution to the population variance of blood pressure has been conclusively identified.

Mendelian forms of hypertension

The renin–angiotensin–aldosterone system (summarized in Figure 25.1) is a central pathway in the maintenance of salt and water balance and blood pressure; salt and water retention consequent on increased mineralocorticoid activity acts via a negative feedback loop to reduce plasma concentrations of renin. Over the last ten years, the genes responsible for a number of rare Mendelian forms of hypertension have been identified in a series of elegant studies that have contributed much to our understanding of the molecular

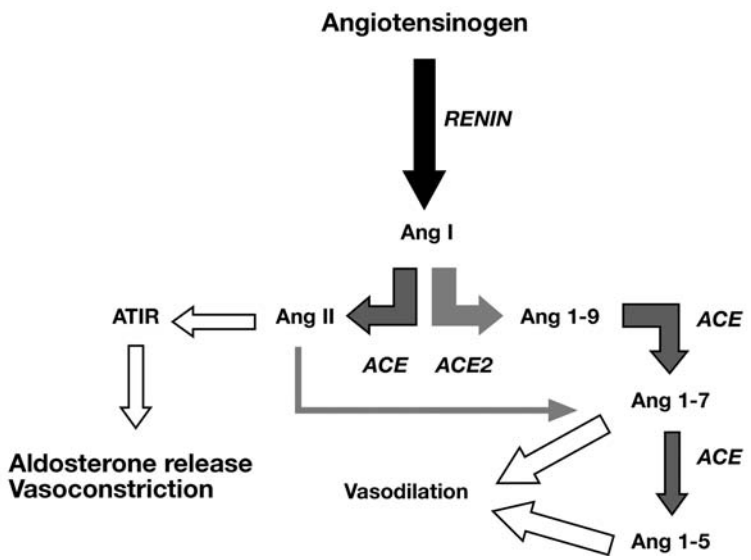

Figure 25.1 The renin-angiotensin-aldosterone system. Angiotensinogen is synthesized by the liver and undergoes enzymatic cleavage by renin released from the kidney to produce the decapeptide angiotensin-I. Angiotensin-I is cleaved by ACE to produce angiotensin-II, which acts via the angiotensin-II type I receptor to promote the release of aldosterone, which increases salt and water retention in the distal nephron. Angiotensin-I is also cleaved by the recently discovered enzyme ACE2 to form Ang (1–9), which is further processed by ACE to the vasodilatory compounds Ang (1–7) and Ang (1–5). Salt and water retention via aldosterone and vasoconstriction due to the action of Ang II increases the kidney glomerular filtration pressure which inhibits the release of renin completing the negative feedback loop.

physiology of blood pressure control. The ''final common pathway'' for these conditions is an increase in sodium reabsorption in the distal nephron owing to a pathological overactivity of one of a series of molecules involved in the mineralocorticoid signaling pathway. Genes causing syndromes involving hyper- or hypo-tension are summarized in Table 25.1. These insights have promoted the conduct of studies in non-Mendelian hypertension, particularly with low renin, to investigate whether common polymorphic variants in these same genes can be susceptibility factors. So far this question remains largely open since sufficiently large studies that have comprehensively assessed all common variation in such genes are relatively few. A full discussion of each individual Mendelian condition is beyond the scope of this review, so here we focus on four illustrative examples.

Glucocorticoid remediable aldosteronism (GRA)

The last stages in the synthesis of the mineralocorticoid aldosterone and the glucocorticoid cortisol take place in the adrenal cortex, and are catalyzed by two highly homologous enzymes which lie adjacent to each other on chromosome 8. *CYP11B1* (11-beta hydroxylase) converts deoxycortisol to cortisol (the principal corticosteroid), a process involving 11-hydroxylation, and *CYP11B2* (aldosterone synthase) converts deoxycorticosterone to aldosterone (the principal mineralocorticoid), a process involving 18-hydroxylation. The *CYP11B1* promoter is responsive to plasma levels of adrenocorticotrophic hormone (ACTH) and thus *CYP11B1* is constitutively active, regulated by negative feedback of cortisol on the hypothalamic/pituitary axis. The *CYP11B2* promoter responds to the activity of the renin–angiotensin system, mediated via the action of angiotensin II on Type I angiotensin II receptors. In some families segregating hypertension, hypokalemia and hyperaldosteronism in an autosomal dominant fashion, an unequal crossover between these genes (Figure 25.2) results in a chimeric gene having the biological activity of *CYP11B2* (thus, the ability to make aldosterone), but under the control of the ACTH responsive element normally controlling *CYP11B1*. Thus, CYP11B2 activity escapes in affected individuals from the control of the renin–angiotensin system and results in hyperaldosteronism. Because aldosterone does not suppress ACTH release, no negative feedback occurs to limit the production of aldosterone from the chimeric gene; however, the increased blood pressure acts to limit the release of renin from the kidney, leading to a low level of plasma renin. Pharmacological administration of glucocorticoids does suppress ACTH levels and as well as switching off the activity of *CYP11B1*, switches off the chimeric gene and resolves the hyperaldosteronism (Lifton *et al.*, 1992) Since not all patients with GRA have plasma levels of potassium below the conventionally determined normal range, several studies subsequently examined hypertensive cohorts of different ethnicities to determine whether undetected GRA was a common cause of essential hypertension; it was not (MacConnachie *et al.*, 1998; McKenzie *et al.*, 2000). Common polymorphisms in the *CYP11B1* and *CYP11B2* genes have been investigated for association with essential hypertension, but for these genes as for almost all others so investigated, the results from different studies are contradictory. However, a replicable association between polymorphisms at this locus and the efficiency of 11-beta hydroxylation does appear to exist; it remains possible that the effect of the locus on 11-beta hydroxylation plays some causative role in the etiology of hypertension (see below).

Apparent mineralocorticoid excess (AME)

Though cortisol is principally a glucocorticoid, the mineralocorticoid receptor exhibits the same affinity for cortisol in vitro as it does for aldosterone. Since the plasma concentration of cortisol is over 100 times higher than that of aldosterone, the discovery that cortisol could bind to the

Table 25.1. Mutations in genes causing Mendelian syndromes involving hyper- or hypo-tension

Elevated blood pressure			Reduced blood pressure		
Syndrome	Gene(s)	Reference	Syndrome	Gene	Reference
Glucocorticoid remediable aldosteronism	*CYP11B1/B2* chimerism	(Lifton *et al.*, 1992)	Gitelman's variant of Bartter's syndrome	Thiazide-sensitive Na/Cl cotransporter (*ENCC1*)	(Simon *et al.*, 1996*c*, Simon *et al.*, 1997)
Apparent mineralocorticoid excess	*HSD11B2* inactivating mutations	(Mune *et al.*, 1995)	Bartter's syndrome	Renal chloride channel (*CLCNKB*) Na-K-2Cl transporter (*NKCC2*) ATP-sensitive K^+ channel (*ROMK*)	Simon *et al.*, 1997) (Simon *et al.*, 1996*a*) (Simon *et al.*, 1996*b*)
Progesterone-sensitive hypertension	Constitutively active MR	(Geller *et al.*, 2000)	Autosomal dominant pseudohypo-aldosteronism Type I	Loss of function mutations in MR	(Geller *et al.*, 1998)
Liddle's syndrome	Mutations of *SCNN1B* or *SCNN1G*	(Hansson *et al.*, 1995; Shimkets *et al.*, 1994)	Autosomal recessive pseudohypoal-dosteronism Type I	Loss of function mutations in ENaC subunits	(Chang *et al.*, 1996)
Pseudohypoaldoste-ronism type II (Gordon's syndrome)	WNK kinase 1 WNK kinase 4	(Wilson *et al.*, 2001)	Aldosterone synthase deficiency	CYP11B2	(Veldhuis *et al.*, 1980)
Maternally trans-mitted hyperten-sion with hypercholesterole-mia and hypomagnesemia	Mitochondrial *tRNA(Ile)* mutation	(Wilson *et al.*, 2004)	Congenital adrenal hyperplasia secondary to 21-hydroxylase deficiency	CYP21B	(White and New, 1988)
Congenital adrenal hyperplasia secondary to 11-hydroxylase deficiency	*CYP11B1*	(White and New, 1988)			
Congenital adrenal hyperplasia secondary to 17α-hydroxylase deficiency	*CYP17*	(White and New, 1988)			

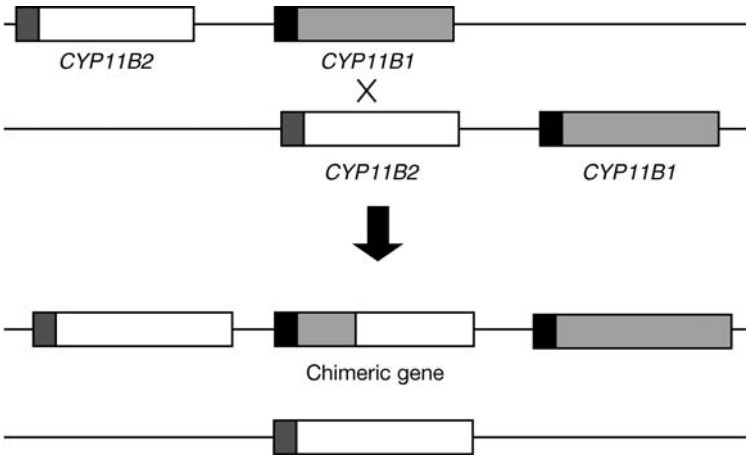

Figure 25.2 Molecular mechanism of GRA. The *CYP11B2* gene (white) has a promoter responsive to the actions of angiotensin-II (green), whereas the *CYP11B1* gene (red) has a promoter responsive to the actions of adrenocorticotrophic hormone from the pituitary (blue). An unequal crossing over event results in the recipient chromosome having intact *CYP11B2* and *B1* genes and also a chimeric gene containing enough *CYP11B2* nucleotides to have aldosterone synthesis capacity, but under the control of the *CYP11B1* promoter. Thus, aldosterone synthesis in people with such a chromosome is at a much higher level in response to the background levels of ACTH, and can be switched off by the administration of glucocorticoids, which inhibit ACTH release.

mineralocorticoid receptor raised the question of how the specificity of that receptor for aldosterone is preserved. This is the role of the short chain dehydrogenase-reductase enzyme 11-beta hydroxysteroid dehydrogenase type 2 (*HSD11B2*), which is principally expressed in the collecting ducts of the kidney and is responsible for converting cortisol to cortisone (which is inactive at the mineralocorticoid receptor [Figure 25.3]). The syndrome of hypertension and hypokalemic alkalosis produced by liquorice intoxication or the drug carbenoxolone (which is very similar biochemically to hyperaldosteronism, but without aldosterone excess) is due to the pharmacological inhibition of *HSD11B2* and subsequent exposure of the mineralocorticoid receptor to the actions of cortisol. AME is a rare recessive syndrome (less than 100 families in total having been reported), characterized by childhood onset of hypertension with hypokalemic alkalosis and low plasma renin and aldosterone levels; it was hypothesized that AME was also due to a deficiency in 11-beta-hydroxysteroid dehydrogenase activity. The *HSD11B2* gene is located on chromosome 16q22, is 6.2 kb in length, and contains five exons. In 1995 mutations in both alleles of *HSD11B2* in eight kindreds containing eleven affected individuals with AME were discovered; these mutations markedly reduced enzymatic activity and enabled cortisol to occupy the mineralocorticoid receptor, producing the observed phenotype and low levels of both plasma renin and aldosterone (Mune *et al.*, 1995). Over 30 different mutations have since been defined within the *HSD11B2* gene in approximately 60 affected kindreds (Quinkler and Stewart, 2003). With respect to susceptibility to essential hypertension, a microsatellite in the first intron of *HSD11B2* has been evaluated for linkage and association with negative results, however, shorter alleles of this microsatellite have been found to be associated with salt sensitivity in studies involving two ethnic groups. Thus far, studies of sufficient size to test this hypothesized gene–environment interaction robustly are lacking.

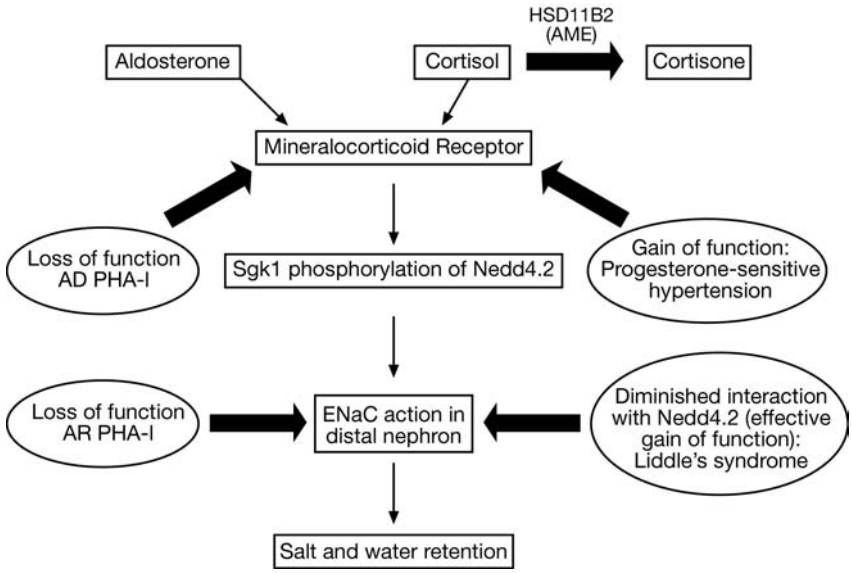

Figure 25.3 Mineralocorticoid signaling and hypertension. Conditions causing disturbances of blood pressure are indicated by blue arrows. AD-PHA I: autosomal dominant pseudohypoaldosteronism type I; AR-PHA I: autosomal recessive pseudohypoaldosteronism type I. Increased Sgk-1 activity consequent on mineralocorticoid receptor activation phosphorylates NEDD 4–2, which diminishes ENaC ubiquitination and thus increases its activity at the tubular cell surface.

Hypertension and hypotension due to MR mutations

Heterozygous loss-of-function mutations in the mineralocorticoid receptor (MR) cause autosomal dominant pseudohypoaldosteronism Type I, which is characterized by salt wasting and hypotension (Figure 25.3) (Geller *et al.*, 1998). Under physiological conditions, activation of the MR by aldosterone causes an increase in distal tubular renal sodium reabsorption by the epithelial sodium channel of the distal nephron. Geller and colleagues screened the *MR* gene for mutations in 75 patients with an early onset of severe hypertension to test the hypothesis that gain-of-function mutations in the *MR* gene could cause hypertension. They found that one teenage patient with severe hypertension, suppressed plasma renin activity, low serum aldosterone, and no other underlying cause of hypertension was heterozygous for a missense mutation, resulting in substitution of leucine for serine at codon 810 in the *MR* gene. The S810L mutation was found to lie in the MR hormone-binding domain (HBD), altering an amino acid that is conserved in all MRs from *Xenopus* to human. Family screening of the proband revealed Mendelian segregation of early onset hypertension with the MR mutation. Functional studies revealed not only constitutive activity of the mutant MR, but also that the mutant protein was activated by a wider range of steroid hormones than the wild-type protein, particularly progesterone. Progesterone levels increase by over a hundredfold in pregnancy, and indeed in female members of this kindred carrying the mutation, severe exacerbations of hypertension in pregnancy were reported (Geller *et al.*, 2000).

Liddle's syndrome

The epithelial sodium channel (ENaC) is expressed in the cortical and medullary collecting tubule of

the nephron and is responsible for sodium reabsorption (and, through the electrochemical gradient generated across the renal epithelial cells, for potassium secretion). The channel has three subunits (α, β and γ). ENaC is the final effector molecule for the mineralocorticoid signaling pathway (Figure 25.3). Aldosterone signaling via the MR increases the intracellular activity of the serine/threonine kinase Sgk1, which phosphorylates the ubiquitin-protein ligase Nedd 4–2. Phosphorylation of Nedd 4–2 diminishes its interaction with ENaC and thus decreases ENaC ubiquitination. This increases the amount of ENaC expressed on renal tubular cell membranes by downregulating either ubiquitin-dependent receptor endocytosis or ubiquitin-dependent lysosomal receptor degradation. ENaC activity is inhibited by amiloride, a mild diuretic generally ineffective as an antihypertensive when used alone. Homozygous mutations causing loss-of-function in ENaC subunits lead to a severe salt-wasting condition biochemically similar to hypoaldosteronism, but with supranormal aldosterone levels (autosomal recessive pseudohypoaldosteronism type I; Figure 25.3) (Chang *et al.*, 1996). By contrast, gain of function mutations in the β or γ subunit lead to increased sodium and water retention (and thus hypertension), and increased potassium secretion into the tubule (and thus hypokalemia): such mutations typically present as an autosomal dominant form of hypertension known as Liddle's syndrome which is characterized, as would be expected based on the pathophysiology, by an excellent clinical response to salt restriction or amiloride therapy (Hansson *et al.*, 1995; Shimkets *et al.*, 1994). Mutations causing Liddle's syndrome are found in the carboxy-terminal ends of the β or γ subunits and disrupt an amino acid motif known as ''PY'', which consists of the amino acids PPXY (where P represents proline, Y represents tyrosine, and X represents any amino acid). The PY motif interacts with a tryptophan–tryptophan (WW) motif on Nedd4–2; an intact PY motif is required for this interaction.

Although Mendelian forms of hypertension have yielded much information about the regulation of salt and water homeostasis and blood pressure regulation, investigation of these genes as candidates for essential hypertension in association studies has not, so far, yielded strong evidence for their involvement. It is, however, clear that for a multifactorial disease of relatively low heritability such as hypertension, studies of many thousands of cases will be necessary to establish robustly the role of particular polymorphisms; thus far, such studies are lacking. The information gained from the Mendelian forms of hypertension strongly suggests that essential hypertension will be a highly genetically heterogeneous disorder, which is likely to magnify even further the size of studies that will be required.

Genetics of blood pressure and essential hypertension

A number of genome-wide genetic linkage studies of hypertension in affected relative pairs (Caulfield *et al.*, 2003; Kardia *et al.*, 2003; Rao *et al.*, 2003; von Wowern *et al.*, 2003), quantitative levels of blood pressure (de Lange *et al.*, 2004; James *et al.*, 2003) or other study designs have been undertaken in an attempt to localize genes with substantial effects. Candidate gene investigations comparing case/control allele frequencies have also been widely applied in studies of hypertension.

Linkage studies

Two features that are critical for the design of linkage studies are the diagnostic criteria used for inclusion of patients, and the sample size (number of affected sib pairs or other informative relative pairs). Statistical evaluation under simple genetic models show that adequate power to detect loci that contribute say ~5% of blood pressure variance will only be obtained in linkage studies with more than 1000 affected relative pairs selected in the upper extreme of the blood pressure distribution

(e.g. the upper 5% or even better, upper 1% of the blood pressure distribution adjusted for age and sex). Only a few linkage investigations to date could be said to approach these power requirements, notably the British Genetics of Hypertension Study (BRIGHT) (Caulfield *et al.*, 2003) and the NIH HyperGen study (Rao *et al.*, 2003). BRIGHT, which is the largest individual study, comprises more than 2000 affected sibpairs selected to be in the upper 10%–5% of the blood pressure distribution. Based on genome-wide linkage mapping, four regions of suggestive linkage were reported in BRIGHT, including one region with a LOD score of 3.2. Different regions with suggestive evidence of linkage to hypertension susceptibility have been observed in the HyperGen study and the other investigations cited above.

A feature in many of these genetic investigations has been the incorporation of additional traits that are related to hypertension. For example, HyperGen has included linkage studies of pulse pressure, echocardiography, heart rate, lipids, age-at-diagnosis and other phenotypes measured in the hypertensive families, but again only suggestive evidence of linkage have been reported. The analysis of blood pressure measurements as quantitative traits has revealed stronger evidence of linkage in some studies, particularly the Framingham Heart Study for which longitudinal data can be analysed (James *et al.*, 2003; Levy *et al.*, 2000). Some other intermediate phenotypes for blood pressure are discussed below.

The linkage results for hypertension are similar to those seen in many other multifactorial diseases. They suggest that the loci underlying susceptibility to hypertension have at best a modest effect in the populations from which these study samples have been selected. Nevertheless, it may be possible to identify some of the susceptibility genes by a combination of additional linkage and association mapping. Further mapping and additional meta-analyses involving the pooling of data from different genome screens may be helpful in confirming or identifying new linkage regions (Liu *et al.*, 2004; Province *et al.*, 2003).

Candidate gene investigations

The literature on candidate genes investigated for association with hypertension is vast, but the results are often inconclusive. Angiotensinogen (AGT), which is amongst the most extensively studied candidate genes for hypertension, illustrates the issues. Since the first report of AGT polymorphisms associated with hypertension (Jeunemaitre *et al.*, 1992), many similar studies have been undertaken. More than 30 published association studies are available on this candidate gene just for the period from 2003 through early 2005. Many of these report positive evidence of association with hypertension or blood pressure, but most involve small cohorts of patients and/or a subset with specific clinical features (e.g. high sodium intake (Yamagishi *et al.*, 2004), response to therapy (Kurland *et al.*, 2004), left-ventricular mass (Kuznetsova *et al.*, 2005)). Amongst the larger studies during this period, Ortlepp *et al.* (Ortlepp *et al.*, 2003) have reported significantly increased diastolic blood pressure associated with the AGT *M235T* site T allele in a German cohort of 1903 young men, whereas Renner *et al.* (Renner *et al.*, 2005) found no relationship of AGT polymorphisms with blood pressure in a second German cohort of 2575 individuals most of whom had a history of cardiovascular disease. In a study of 212 untreated hypertensive patients in Germany, Brand-Herrmann *et al.* (Brand-Herrmann *et al.*, 2004) found association of AGT haplotypes with blood pressure, and Wu *et al.* (Wu *et al.*, 2004) have also described a significant association of AGT haplotypes with the study of 461 hypertensive patients and 327 normotensive controls from Taiwan. Matsubara *et al.* (Matsubara *et al.*, 2003) report a weak association of AGT *M235T* with pulse pressure in a community survey of 1245 Japanese over 40 years of age.

These results are illustrative of the difficulty of evaluating and comparing results for candidate gene evaluation in the literature. The situation is unlikely to be clarified until systematic evaluation is undertaken with standardised methodologies in

studies involving very large (e.g. several thousand patients) cohorts with well-defined clinical characteristics. Meta-analyses in which the results from different studies are combined may be useful for hypertension, as has proved to be the case in other disease areas, e.g. calpain-10 variants in non-insulin-dependent diabetes mellitus (Song *et al.*, 2004; Weedon *et al.*, 2003) or the angiotensin-converting enzyme (*ACE*) gene and myocardial infarction (Keavney *et al.*, 2000); meta-analysis may also help detect publication bias leading to the over-representation of positive results in small samples (Keavney *et al.*, 2000). However, meta-analysis in the area of hypertension research will require very careful attention to patient inclusion criteria as these are particularly variable in this field.

An interesting source of candidate genes for hypertension is arising through the investigation of genetic models of blood regulation in rodents, particularly the rat (Stoll and Jacob, 2001). As an example, we cite recent work on the type II SH2 domain containing the inositol 5-phosphatase (*SHIP2*) gene (Clement *et al.*, 2001). *SHIP2* maps to a region of linkage to traits related to metabolic syndrome on rat chromosome 1 in several disease models including Goto-Kakizaki (GK) diabetic rat and spontaneously hypertensive rat strains. These two strains harbour an identical *SHIP2* mutation that disrupts a potential class II ligand for Src homology (SH)-3 domain and slightly impairs insulin signaling in vitro (Marion *et al.*, 2002). Analysis of *SHIP2* SNPs in UK patients with non-insulin dependent diabetes and their siblings showed strong evidence of association with hypertension ($p = 9.3 \times 10^{-6}$) and related traits; in French diabetics, an association was found only with hypertension ($p = 0.01$) (Kaisaki *et al.*, 2004). Although the association of the *SHIP2* variants with hypertension is very strong in at least one cohort, in contrast with many other candidate gene investigations, nevertheless, the examination of the gene in additional cohorts of hypertensive patients selected through the presence of diabetes is required as further confirmation of these results.

Intermediate phenotypes

The study of intermediate phenotypes for hypertension is attractive for a number of reasons. Firstly, such phenotypes may be under a greater degree of genetic control than the final phenotype, and thus studies of moderate size may suffice to detect genetic effects. Secondly, study of an intermediate phenotype may enable the blood pressure phenotype to be subdivided into more genetically tractable subphenotypes. Thirdly, in order to study these phenotypes, ascertainment from the extremes of the population with its concomitant practical difficulties may not be required. Fourthly, many compounds that can be measured in blood have smaller measurement errors than can be achieved for blood pressure in very large numbers of subjects. By contrast with the hypertension phenotype, robust associations between markers in some genes and particular traits have been described; as yet, however, the relevance of these traits to hypertension has not been shown. We discuss two such robust associations: between the angiotensin-converting enzyme (*ACE*) gene and plasma levels of *ACE*, and between the *CYP11B1/B2* locus and steroid 11-hydroxylation.

The *ACE* gene and trait

Given the pivotal role of ACE in the renin–angiotensin system and the demonstrated efficacy of ACE inhibitors in the treatment of hypertension, the *ACE* gene and trait have understandably been among the most studied with respect to hypertension and cardiovascular disease. Plasma levels of *ACE* are under very strong genetic control, with about 70% of the total variability being determined by genes. Moreover, complex segregation analysis showed that around two-thirds of this heritability could be ascribed to a single locus, subsequently identified as the *ACE* gene itself on chromosome 17 (Rigat *et al.*, 1990; Tiret *et al.*, 1992). The very strong genetic signal at the *ACE* locus enabled fine-scale genetic dissection of the locus to be

undertaken: during the course of these investigations, *ACE* was one of the first loci shown to exhibit the limited haplotype diversity and haplotype ''block'' structure subsequently found throughout the human genome; and the utility of evolutionary and trans-ethnic approaches to fine-map causative variants in a gene was convincingly demonstrated (Farrall *et al.*, 1999; Keavney *et al.*, 1998; McKenzie *et al.*, 2001). A recent study showing that the *ACE* trait is influenced by at least two major intragenic and one minor promoter site variant has provided us with perhaps the most detailed information available with regard to the genetic control of any quantitative trait (Soubrier *et al.*, 2002). However, studies in humans which have looked for association between the polymorphisms associated with plasma *ACE* activity (chiefly, but not exclusively, the insertion/deletion or I/D polymorphism) and hypertension have been mainly negative. Moreover, there is no strong evidence for association between plasma *ACE* levels and blood pressure. Thus, although investigation of plasma *ACE* has provided us with many insights regarding the genetic control of that phenotype, it has not been established as a true intermediate phenotype for hypertension.

Steroid 11-hydroxylation and the *CYP11B1/B2* locus

The function and genomic location of the *CYP11B1* and *CYP11B2* genes are described above. Following the identification of this locus as causative in GRA, polymorphisms principally in the CYP11B2 gene were examined for association with essential hypertension. Results were discrepant, but a more robust association between a 5′ variant of *CYP11B2 (C-344T)* and elevated plasma levels of 11-deoxycortisol was detected in several different studies (Connell *et al.*, 2004; Davies *et al.*, 1999; 2001). 11-deoxycortisol has no biological function but is the immediate precursor of cortisol and the substrate for *CYP11B1*. Studies which examined the urinary steroid profile showed that the principal urinary metabolite of 11-deoxycortisol, tetrahydrodeoxycortisol (THS) was also associated with genotype at *CYP11B2*. Thus, the T allele at C-344T appears to be associated with a lower efficiency of 11-hydroxylation. Since 11-hydroxylation is the function of *CYP11B1* rather than *CYP11B2*, it was hypothesized that the effect of C-344T was via linkage disequilibrium with a causative variant or variants in *CYP11B1*. Recently, it was demonstrated that polymorphisms in *CYP11B1* were stronger predictors of 11-hydroxylation than C-344T; fine mapping of the *CYP11B1* locus may reveal the variants responsible for these observations (Keavney *et al.*, 2005). Some hypertensive patients have evidence of impairment of 11-hydroxylation, and Connell and colleagues have suggested that relative inefficiency of 11-hydroxylation (and thus, relative cortisol deficiency) may upregulate ACTH drive to the adrenal cortex and increase lifelong exposure to ACTH and other POMC-derived peptides, potentially resulting in relative aldosterone excess and predisposing to hypertension in a subset of people (Connell *et al.*, 2003). In support of this hypothesis, hypertensive patients with a raised aldosterone to renin ratio have been shown to have a higher frequency of those alleles predisposing to less efficient 11-hydroxylation. By contrast with ACE, the heritability of steroid metabolite levels is only moderate (for example, 20% for urinary THS excretion) and it may be that identification of the quantitative trait nucleotide(s) at *CYP11B1* may be correspondingly more difficult.

Conclusions and perspectives

As discussed above, investigations of Mendelian forms of hypertension over the last few years have successfully led to the identification of several novel genes involved in these diseases, and yielded important new information on the regulation of salt and water homeostasis and blood pressure. In contrast, the genetics of essential hypertension has proved refractory to molecular genetics approaches. This situation is similar to that for other multifactorial disorders, and is due in part to

inadequacy of genomic tools. With much improved knowledge of the human genome sequence and its gene content following the international sequencing effort, and the development of high-density SNP maps suitable for association studies, this situation has been addressed. We can expect new studies to evaluate systematically genes in regions that have exhibited suggestive evidence of linkage to determine if they harbour variants that are associated with hypertension or related traits. In parallel, a large number of candidate genes can now be investigated with a high density of markers using information on SNPs from databases and high-throughput genotyping methodology. In the near future, association studies will undoubtedly be attempted for these phenotypes with a whole-genome approach, using markers that are intended to capture linkage disequilibrium patterns in all human genes and/or across all of the human genome.

Despite the improvement in genomic tools, the task of identifying the genes underlying essential hypertension in humans remains daunting. Additional genetic mapping studies in specific families or populations chosen to complement existing cohorts are likely to be informative. For example, autosomal dominant hypertension in a large family of Turkish origin has been successfully mapped to a small region of chromosome 12 harboring complex chromosome deletions and rearrangments, leading to the identification of a small number of potential candidate genes (Bahring *et al.*, 2004). Reduced genetic complexity of hypertension and blood pressure regulation may also be found in genetically isolated populations. Populations with a high degree of consanguinity may be of interest for mapping loci involved in blood pressure if many of these are recessively acting, as suggested by quantitative genetic analysis (Rudan *et al.*, 2003). Comparative studies in different species are likely to have a key role in genetic studies of hypertension and blood pressure regulation over the next few years given the increasing effort to identify positional candidate genes in rodent models.

Considerable attention is needed to improve study design for evaluation of the candidate genes that will emerge from such investigations. The field of hypertension genetics would benefit greatly if a consensus could be developed around a number of large standardized patient and control cohorts for genetic investigations. The selection of any cohort requires careful consideration to ensure that the phenotype is sufficiently extreme to provide adequate power. Unless rigorous selection can be applied to obtain patients in the upper 5%, or better 1%, of the blood pressure distribution, the power of genetic studies is likely to be inadequate even in large cohorts. This is a particular issue for hypertension because the disease is common, has varying severity at diagnosis, and is modified by treatment. Comparisons between studies, meta-analyses and replication of associations will remain difficult until the criteria used to constitute the study populations are clearly defined and strictly applied.

REFERENCES

Bahring, S., Rauch, A., Toka, O. *et al.* (2004). Autosomal-dominant hypertension with type E brachydactyly is caused by rearrangement on the short arm of chromosome 12. *Hypertension*, **43**, 471–6.

Brand-Herrmann, S. M., Kopke, K., Reichenberger, F. *et al.* (2004). Angiotensinogen promoter haplotypes are associated with blood pressure in untreated hypertensives. *J Hypertens*, **22**, 1289–97.

Caulfield, M., Munroe, P., Pembroke, J. *et al.* (2003). Genome-wide mapping of human loci for essential hypertension. *Lancet*, **361**, 2118–23.

Chang, S. S., Grunder, S., Hanukoglu, A. *et al.* (1996). Mutations in subunits of the epithelial sodium channel cause salt wasting with hyperkalaemic acidosis, pseudohypoaldosteronism type 1. *Nat Genet*, **12**, 248–53.

Clement, S., Krause, U., Desmedt, F. *et al.* (2001). The lipid phosphatase SHIP2 controls insulin sensitivity. *Nature*, **409**, 92–7.

Connell, J. M., Fraser, R., MacKenzie, S. and Davies, E. (2003). Is altered adrenal steroid biosynthesis a key

intermediate phenotype in hypertension? *Hypertension*, **41**, 993–9.

Connell, J.M., Fraser, R., MacKenzie, S.M. *et al.* (2004). The impact of polymorphisms in the gene encoding aldosterone synthase (*CYP11B2*) on steroid synthesis and blood pressure regulation. *Mol Cell Endocrinol*, **217**, 243–7.

Davies, E., Holloway, C.D., Ingram, M.C. *et al.* (2001). An influence of variation in the aldosterone synthase gene (*CYP11B2*) on corticosteroid responses to ACTH in normal human subjects. *Clin Endocrinol (Oxf)*, **54**, 813–17.

Davies, E., Holloway, C.D., Ingram, M.C. *et al.* (1999). Aldosterone excretion rate and blood pressure in essential hypertension are related to polymorphic differences in the aldosterone synthase gene *CYP11B2*. *Hypertension*, **33**, 703–7.

de Lange, M., Spector, T.D. and Andrew, T. (2004). Genome-wide scan for blood pressure suggests linkage to chromosome 11, and replication of loci on 16, 17, and 22. *Hypertension*, **44**, 872–7.

European Society of Hypertension/European Society of Cardiology Guidelines Committee. (2003). European Society of Hypertension–European Society of Cardiology guidelines for the management of arterial hypertension. *J Hypertens*, **21**, 1011–53.

Fagard, R., Brguljan, J., Staessen, J. *et al.* (1995). Heritability of conventional and ambulatory blood pressures. A study in twins. *Hypertension*, **26**, 919–24.

Farrall, M. (2004). Quantitative genetic variation: a post-modern view. *Hum Mol Genet*, **13**, Spec No 1, R1–7.

Farrall, M., Keavney, B., McKenzie, C. *et al.* (1999). Fine-mapping of an ancestral recombination breakpoint in DCP1. *Nat Genet*, **23**, 270–1.

Fava, C., Burri, P., Almgren, P. *et al.* (2004). Heritability of ambulatory and office blood pressure phenotypes in Swedish families. *J Hypertens*, **22**, 1717–21.

Geller, D.S., Farhi, A., Pinkerton, N. *et al.* (2000). Activating mineralocorticoid receptor mutation in hypertension exacerbated by pregnancy. *Science*, **289**, 119–23.

Geller, D.S., Rodriguez-Soriano, J., Vallo Boado, A. *et al.* (1998). Mutations in the mineralocorticoid receptor gene cause autosomal dominant pseudohypoaldosteronism type I. *Nat Genet*, **19**, 279–81.

Hansson, J.H., Nelson-Williams, C., Suzuki, H. *et al.* (1995). Hypertension caused by a truncated epithelial sodium channel gamma subunit: genetic heterogeneity of Liddle syndrome. *Nat Genet*, **11**, 76–82.

He, F.J. and MacGregor, G.A. (2003). Cost of poor blood pressure control in the UK: 62,000 unnecessary deaths per year. *J Hum Hypertens*, **17**, 455–7.

James, K., Weitzel, L.R., Engelman, C.D., Zerbe, G. and Norris, J.M. (2003). Genome scan linkage results for longitudinal systolic blood pressure phenotypes in subjects from the Framingham Heart Study. *BMC Genet*, **4** Suppl 1, S83.

Jeunemaitre, X., Soubrier, F., Kotelevtsev, Y.V. *et al.* (1992). Molecular basis of human hypertension: role of angiotensinogen. *Cell*, **71**, 169–80.

Kaisaki, P.J., Delepine, M., Woon, P.Y. *et al.* (2004). Polymorphisms in type II SH2 domain-containing inositol 5-phosphatase (INPPL1, SHIP2) are associated with physiological abnormalities of the metabolic syndrome. *Diabetes*, **53**, 1900–4.

Kardia, S.L., Rozek, L.S., Krushkal, J. *et al.* (2003). Genome-wide linkage analyses for hypertension genes in two ethnically and geographically diverse populations. *Am J Hypertens*, **16**, 154–7.

Keavney, B., Mayosi, B., Gaukrodger, N. *et al.* (2005). Genetic variation at the locus encompassing 11-beta hydroxylase and aldosterone synthase accounts for heritability in cortisol precursor (11-deoxycortisol) urinary metabolite excretion. *J Clin Endocrinol Metab*, **90**, 1072–7.

Keavney, B., McKenzie, C., Parish, S. *et al.* (2000). Large-scale test of hypothesised associations between the angiotensin-converting-enzyme insertion/deletion polymorphism and myocardial infarction in about 5000 cases and 6000 controls. International Studies of Infarct Survival (ISIS) Collaborators. *Lancet*, **355**, 434–42.

Keavney, B., McKenzie, C.A., Connell, J.M. *et al.* (1998). Measured haplotype analysis of the angiotensin-I converting enzyme gene. *Hum Mol Genet*, **7**, 1745–51.

Kupper, N., Willemsen, G., Riese, H. *et al.* (2005). Heritability of daytime ambulatory blood pressure in an extended twin design. *Hypertension*, **45**, 80–5.

Kurland, L., Liljedahl, U., Karlsson, J. *et al.* (2004). Angiotensinogen gene polymorphisms: relationship to blood pressure response to antihypertensive treatment. Results from the Swedish Irbesartan Left Ventricular Hypertrophy Investigation vs Atenolol (SILVHIA) trial. *Am J Hypertens*, **17**, 8–13.

Kuznetsova, T., Staessen, J.A., Reineke, T. *et al.* (2005). Context-dependency of the relation between left ventricular mass and AGT gene variants. *J Hum Hypertens*, **19**, 155–63.

Levy, D., DeStefano, A.L., Larson, M.G. *et al.* (2000). Evidence for a gene influencing blood pressure on chromosome 17. Genome scan linkage results for longitudinal blood pressure phenotypes in subjects from the framingham heart study. *Hypertension*, **36**, 477–83.

Lewington, S., Clarke, R., Qizilbash, N., Peto, R. and Collins, R. (2002). Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet*, **360**, 1903–13.

Lifton, R.P., Dluhy, R.G., Powers, M. *et al.* (1992). A chimaeric 11 beta-hydroxylase/aldosterone synthase gene causes glucocorticoid-remediable aldosteronism and human hypertension. *Nature*, **355**, 262–5.

Liu, W., Zhao, W. and Chase, G.A. (2004). Genome scan meta-analysis for hypertension. *Am J Hypertens*, **17**, 1100–6.

MacConnachie, A.A., Kelly, K.F., McNamara, A. *et al.* (1998). Rapid diagnosis and identification of cross-over sites in patients with glucocorticoid remediable aldosteronism. *J Clin Endocrinol Metab*, **83**, 4328–31.

Marion, E., Kaisaki, P.J., Pouillon, V. *et al.* (2002). The gene *INPPL1*, encoding the lipid phosphatase SHIP2, is a candidate for type 2 diabetes in rat and man. *Diabetes*, **51**, 2012–17.

Matsubara, M., Metoki, H., Katsuya, T. *et al.* (2003). T+31C polymorphism (M235T) of the angiotensinogen gene and home blood pressure in the Japanese general population: the Ohasama Study. *Hypertens Res*, **26**, 47–52.

McKenzie, C.A., Abecasis, G.R., Keavney, B. *et al.* (2001). Trans-ethnic fine mapping of a quantitative trait locus for circulating angiotensin I-converting enzyme (ACE). *Hum Mol Genet*, **10**, 1077–84.

McKenzie, C.A., Keavney, B., Forrester, T., Julier, C. and Ratcliffe, P.J. (2000). Screening for the GRA mutation in Jamaica. *J Hum Hypertens*, **14**, 157–8.

Mune, T., Rogerson, F.M., Nikkila, H., Agarwal, A.K. and White, P.C. (1995). Human hypertension caused by mutations in the kidney isozyme of 11 beta-hydroxysteroid dehydrogenase. *Nat Genet*, **10**, 394–9.

Ortlepp, J.R., Metrikat, J., Mevissen, V. *et al.* (2003). Relation between the angiotensinogen (*AGT*) *M235T* gene polymorphism and blood pressure in a large, homogeneous study population. *J Hum Hypertens*, **17**, 555–9.

Province, M.A., Kardia, S.L., Ranade, K. *et al.* (2003). A meta-analysis of genome-wide linkage scans for hypertension: the National Heart, Lung and Blood Institute Family Blood Pressure Program. *Am J Hypertens*, **16**, 144–7.

Quinkler, M. and Stewart, P.M. (2003). Hypertension and the cortisol-cortisone shuttle. *J Clin Endocrinol Metab*, **88**, 2384–92.

Rao, D.C., Province, M.A., Leppert, M.F. *et al.* (2003). A genome-wide affected sibpair linkage analysis of hypertension: the HyperGen network. *Am J Hypertens*, **16**, 148–50.

Renner, W., Nauck, M., Winkelmann, B.R. *et al.* (2005). Association of angiotensinogen haplotypes with angiotensinogen levels but not with blood pressure or coronary artery disease: the Ludwigshafen Risk and Cardiovascular Health Study. *J Mol Med*, **83**, 235–9.

Rigat, B., Hubert, C., Alhenc-Gelas, F. *et al.* (1990). An insertion/deletion polymorphism in the angiotensin I-converting enzyme gene accounting for half the variance of serum enzyme levels. *J Clin Invest*, **86**, 1343–6.

Rudan, I., Rudan, D., Campbell, H. *et al.* (2003). Inbreeding and risk of late onset complex disease. *J Med Genet*, **40**, 925–32.

Shimkets, R.A., Warnock, D.G., Bositis, C.M. *et al.* (1994). Liddle's syndrome: heritable human hypertension caused by mutations in the beta subunit of the epithelial sodium channel. *Cell*, **79**, 407–14.

Simon, D.B., Bindra, R.S., Mansfield, T.A. *et al.* (1997). Mutations in the chloride channel gene, *CLCNKB*, cause Bartter's syndrome type III. *Nat Genet*, **17**, 171–8.

Simon, D.B., Karet, F.E., Hamdan, J.M. *et al.* (1996*a*). Bartter's syndrome, hypokalaemic alkalosis with hypercalciuria, is caused by mutations in the Na-K-2Cl cotransporter NKCC2. *Nat Genet*, **13**, 183–8.

Simon, D.B., Karet, F.E., Rodriguez-Soriano, J. *et al.* (1996*b*). Genetic heterogeneity of Bartter's syndrome revealed by mutations in the K+ channel, ROMK. *Nat Genet*, **14**, 152–6.

Simon, D.B., Nelson-Williams, C., Bia, M.J. *et al.* (1996*c*). Gitelman's variant of Bartter's syndrome, inherited hypokalaemic alkalosis, is caused by mutations in the thiazide-sensitive Na–Cl cotransporter. *Nat Genet*, **12**, 24–30.

Song, Y., Niu, T., Manson, J.E., Kwiatkowski, D.J. and Liu, S. (2004). Are variants in the *CAPN10* gene related to risk of type 2 diabetes? A quantitative assessment of population and family-based association studies. *Am J Hum Genet*, **74**, 208–22.

Soubrier, F., Martin, S., Alonso, A. *et al.* (2002). High-resolution genetic mapping of the ACE-linked QTL influencing circulating ACE activity. *Eur J Hum Genet*, **10**, 553–61.

Stoll, M. and Jacob, H.J. (2001). Genetic rat models of hypertension: relationship to human hypertension. *Curr Hypertens Rep*, **3**, 157–64.

Tiret, L., Rigat, B., Visvikis, S., Breda, C. *et al.* (1992). Evidence, from combined segregation and linkage analysis, that a variant of the angiotensin I-converting enzyme (ACE) gene controls plasma ACE levels. *Am J Hum Genet*, **51**, 197–205.

Veldhuis, J.D., Kulin, H.E., Santen, R.J., Wilson, T.E. and Melby, J.C. (1980). Inborn error in the terminal step of aldosterone biosynthesis. Corticosterone methyl oxidase tpe II deficiency in a North American pedigree. *N Engl J Med*, **303**, 117–21.

Von Wowern, F., Bengtsson, K., Lindgren, C.M. *et al.* (2003). A genome wide scan for early onset primary hypertension in Scandinavians. *Hum Mol Genet*, **12**, 2077–81.

Weedon, M.N., Schwarz, P.E., Horikawa, Y. *et al.* (2003). Meta-analysis and a large association study confirm a role for calpain-10 variation in type 2 diabetes susceptibility. *Am J Hum Genet*, **73**, 1208–12.

White, P.C. and New, M.I. (1988). Molecular genetics of congenital adrenal hyperplasia. *Baillieres Clin Endocrinol Metab*, **2**, 941–65.

Wilson, F.H., Disse-Nicodeme, S., Choate, K.A. *et al.* (2001). Human hypertension caused by mutations in WNK kinases. *Science*, **293**, 1107–12.

Wilson, F.H., Hariri, A., Farhi, A. *et al.* (2004). A cluster of metabolic defects caused by mutation in a mitochondrial tRNA. *Science*, **306**, 1190–4.

Wu, S.J., Chiang, F.T., Chen, W.J. *et al.* (2004). Three single-nucleotide polymorphisms of the angiotensinogen gene and susceptibility to hypertension: single locus genotype vs. haplotype analysis. *Physiol Genomics*, **17**, 79–86.

Yamagishi, K., Iso, H., Tanigawa, T. *et al.* (2004). High sodium intake strengthens the association between angiotensinogen T174M polymorphism and blood pressure levels among lean men and women: a community-based study. *Hypertens Res*, **27**, 53–60.

Yusuf, S., Hawken, S., Ounpuu, S. *et al.* (2004). Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet*, **364**, 937–52.

26

Obstructive pulmonary disease

Bipen D. Patel and David A. Lomas

Introduction

Obstructive pulmonary disease describes a heterogeneous group of diseases which are characterized by airflow obstruction on expiration. This is defined as a forced expiratory volume in one second (FEV_1) of less than 80% predicted with a ratio of forced expiratory volume in one second to forced vital capacity (FEV_1/FVC) of less than 70%. Asthma and chronic obstructive pulmonary disease (COPD) account for the vast majority of the global disease burden from obstructive pulmonary disease and these two conditions will be the focus of this chapter.

Asthma

Asthma is an inflammatory condition of the airways which results in bronchial reactivity to a variety of stimuli with consequent narrowing of the airways. This results in airflow obstruction which is usually reversible (National Heart, Lung, and Blood Institute, 1992; Department of Health and Social Services, 2003). The cardinal features of airway inflammation, bronchial hyper-reactivity and reversible airflow obstruction give rise to the symptoms of wheeze, breathlessness, cough and chest tightness which are usually the basis for the diagnosis of asthma. Although these symptoms are not specific, the hallmark for the diagnosis of asthma is their intermittent nature, reversibility, diurnal variation and the recognition of specific triggers (British Thoracic Society; Scottish Intercollegiate Guidelines Network, 2003). Asthma is strongly associated with atopy as demonstrated by the link with eczema and hay fever and the detection of IgE, or a positive skin prick test, to a specific allergen (Witt *et al.*, 1986; Woolcock *et al.*, 1987). Asthma is termed extrinsic if it is associated with atopy and intrinsic if it occurs in the absence of atopy.

Environmental risk factors for the development of asthma

The prevalence of asthma has increased in Western countries over the past three to four decades (Anderson *et al.*, 1994; Burney *et al.*, 1990; Peat *et al.*, 1994). In 1995/6 18–21% of children under 15 years of age, and 11% of adults, in England had been diagnosed as having asthma at some point in their life (British Thoracic Society, 2003). The marked increase in asthma strongly suggests that environmental exposures must be instrumental in the development of disease. This is supported by reports in the developing world of an increase in asthma with increasing affluence or urbanization (Keeley *et al.*, 1991; Yemaneberhan *et al.*, 1997) and it has been suggested that a Western lifestyle may predispose individuals to the development of asthma (Woolcock, 1996). A large number of environmental exposures have been identified as risk factors for the development of asthma/atopy (Table 26.1). Three of these, air pollution, childhood infections and diet, will be considered in more detail.

Table 26.1. Environmental risk factors associated with the development of allergic disease

Paternal socioeconomic class
Maternal age at birth
Method of feeding in infancy*
Family size/childhood infection
Childhood vaccination*
Increased exposure to indoor allergens*
Exposure to passive smoking*
Pet ownership
Living on a farm in childhood
Diet*
Increased outdoor pollution*
Change in gut flora
Reduced parasite load*
Population of area
Reduced prevalence of tuberculosis*
Factors associated with enhanced fetal growth*
Increased antibiotic usage*

*Suggested to explain the increased prevalence of asthma.

The role of air pollution in the pathogenesis of asthma is controversial. Increased levels of air pollutants have been reported to increase symptoms in children with established respiratory disease (Roemer *et al.*, 1993). However the prevalence of asthma was lower in children in Leipzig in East Germany than in Munich in West Germany, despite Leipzig having significantly higher levels of air pollution (sulphur dioxide and particulates) (von Mutius *et al.*, 1992). Indeed it has been suggested that the temporal changes in air pollution over recent decades are insufficient to account for the increase in asthma (Seaton *et al.*, 1994).

In 1989, Strachen reported an inverse association between family size and atopy. He suggested that early life infections, transmitted by unhygienic contact with elder siblings, may be protective against the development of allergic disease (Strachan, 1989). This become known as the ''hygiene hypothesis'' and the finding was confirmed in other populations (Bodner *et al.*, 1998; Jarvis *et al.*, 1997; Matricardi *et al.*, 1998; Mattes *et al.*, 1998; Romagnani, 1992; Rona *et al.*, 1997; Strachan *et al.*, 1997*a*; von Mutius *et al.*, 1994). It received support from animal experiments which showed that infections promoted the development of Th1 rather than Th2 lymphocytes. It is the Th2 lymphocytes that are implicated in the development of IgE-mediated immunity and asthma (Romagnani, 1992). There are however a number of limitations to the ''hygiene hypothesis'' as an explanation of the association between family size and asthma/atopy. If the hypothesis were correct, one would expect a significant effect of birth order with the youngest sib afforded greater protection. However several studies have not found this to be the case (Jarvis *et al.*, 1997; Strachan *et al.*, 1997*b*; Svanes *et al.*, 1999). Moreover the effect of early life infection on T cell differentiation is unproven in humans. Even if the hygiene hypothesis is correct, the reduction in family size that has occurred over recent years would only account for about 1% of the observed increase in the prevalence of asthma (Wickens *et al.*, 1999). Thus while smaller family size appears to be associated with an increased risk of allergic disease, it is cannot be concluded that this is the result of decreased childhood infection. Further work is required to understand the role of potential confounding factors in this association such as maternal age, social class, parental education and method of infant feeding.

The temporal increase in asthma prevalence in the UK is at a time when there has been a marked change in diet. The reported association between fruit consumption and respiratory symptoms (Forastiere *et al.*, 2000; Miedema *et al.*, 1993) suggests that dietary anti-oxidants may modify the risk of respiratory disease in susceptible individuals. Dietary intake of vitamin C has been associated with the risk of asthma (Bodner *et al.*, 1999; Schwartz and Weiss, 1990) and vitamin C supplementation may attenuate short and medium term changes in lung function associated with exposure to the oxidant ozone (Grievink *et al.*, 1999; Romieu *et al.*, 1998). Several other anti-oxidant vitamins and co-factors have also been associated with a reduced risk of asthma: vitamin E (Bodner *et al.*, 1999; Hijazi *et al.*, 2000; Troisi *et al.*, 1995), β-carotene (Rautalahti *et al.*, 1997; Troisi *et al.*,

1995), zinc (Schwartz and Weiss, 1990; Soutar *et al.*, 1997), manganese (Soutar *et al.*, 1997), copper (Schwartz and Weiss, 1990) and selenium (Flatt *et al.*, 1990; Stone *et al.*, 1989). However nutrients that are not known to have anti-oxidant activity: magnesium (Britton *et al.*, 1994; Soutar *et al.*, 1997), calcium (Hijazi *et al.*, 2000), niacin (Schwartz and Weiss, 1990) and fatty acids (Bodner *et al.*, 1999; Hodge *et al.*, 1996), have also been associated with a reduced risk of asthma (Fogarty and Britton, 2000). It seems unlikely that each of these nutrients is independently associated with the risk of asthma and further work is required to clarify the association between diet and asthma.

Genetic risk factors for the development of asthma

Twin studies have a shown a greater concordence of asthma in monozygotic than dizygotic twins (Edfors-Lubs, 1971) and more recently family studies have shown greater heritability of extrinsic than intrinsic asthma (The European Community Respiratory Health Survey Group, 1997). As well as confirming a genetic susceptibility to asthma and allergy, these studies showed that inheritance does not follow a mendelian pattern. Linkage studies have identified many genomic regions that may be linked with the asthma and the atopic phenotype. Indeed if all of the loci reported to be associated, at least vaguely, with asthma or atopy were taken into consideration they would constitute about half of the genome (Heinzmann and Deichmann, 2001). Examples of loci that have been reported to be associated with asthma by several different investigators are shown in Table 26.2. In some cases these regions harbour potentially important candidate genes (Heinzmann and Deichmann, 2001). Similarly, association studies have identified a large number of alleles that may be associated with asthma including polymorphisms in *IL-13 (5q-31)*, *CD14 (5q 31)*, *FCER 1β(11q–13)* and *IL-4R (16p12.1)* (Heinzmann and Deichmann, 2001).

Table 26.2. Chromosomal regions reported to associated with asthma/atopy in more than one study

Locus	Candidate genes
5q31–33	*IL-4, IL-13, βADR*
6q21	*TNF-α*
11q-13	*FceR1-β*
12q	
13q	
16p12	*IL-4Rα*

Table 26.3. Examples of loci that have been linked with intermediate phenotypes of asthma

Quantitative phenotype	Locus
Total IgE	*11q13, 7q, 5q, 12q,*
Bronchial hyper-reactivity	*5p 19q*
Blood eosinophil count	*12q24*
Skin prick reactivity	*17q12–21, 16q*
House dust mite sensitivity	*5q31, 8p, 4q, 18p*

The accurate identification of the asthma phenotype is essential for genetic analysis but is not straightforward. The genetic epidemiology of asthma is complicated by the absence of a universally accepted definition and the absence of a specific quantitative marker of asthma. Moreover, asthma may develop at any point in life, remit and recur, with normal health in between (Burney, 1992) and exogenous factors may be less important with increasing age. Thus an alternative strategy is to assess the genetic association with discrete components of the asthma phenotype. Bronchial hyper-reactivity, blood eosinophil count, serum IgE levels and skin prick reactivity have provided quantitative intermediate phenotypes for the study of this complex disease. Table 26.3 shows examples of loci reported to be associated with these intermediate phenotypes (Heinzmann and Deichmann, 2001; Palmer, 2001).

These multiple associations reflect the complexity of asthma genetics. Asthma is a complex genetic disease that is influenced both by the

inflammatory response within the lung and the response of the immune system to environmental exposures. This is exemplified by the recent discovery of two genes, *ADAM-33* and *SPINK5* that have been implicated in conferring genetic susceptibility to asthma by very different molecular mechanisms. The *ADAM-33* gene was identified in a genome wide scan performed on 460 affected sib pairs with asthma. A region on the short arm of chromosome 20 (20p13) showed linkage with asthma, with a maximum log of odds (LOD) score (MLS) of 2.94 (Van Eerdewegh *et al.*, 2002). Despite reducing the numbers by half, the strength of the association increased when the phenotype was refined to current diagnosis of asthma with demonstrable bronchial hyper-reactivity (MLS 3.93). The authors assessed 135 polymorphisms in 23 genes and identified the *ADAM-33* gene as being significantly associated with asthma. The *ADAM* gene family are a subfamily of zinc-dependent metalloproteinases (Shapiro and Owen, 2002). Little is known about the precise role of *ADAM-33* in the development of asthma, and it is not known whether the increase in asthma risk associated with *ADAM-33* is a consequence of increased or reduced synthesis of the active gene product. It is speculated that *ADAM-33* may be associated with fibroblast proliferation, bronchial hyper-reactivity and airway remodeling (Shapiro and Owen, 2002). Thus the increase in asthma risk associated with abnormalities in the activity of *ADAM-33* may be a consequence of alteration in airway function and not through an immunological mechanism.

Walley and colleagues identified polymorphisms of the *SPINK5* gene (5q32) to be associated with the risk of atopy in a study of the families of children with Netherton disease, an autosomal recessive skin disorder that is invariably associated with atopy (Chavanas *et al.*, 2000; Walley *et al.*, 2001). This study found a significant association of the *Glu420Lys* and *Asn368Ser* amino acid substitutions with an increased risk of atopy. There appeared to be parent of origin effect with maternal inheritance of these polymorphisms associated with an approximately four-fold increase in risk of atopy. The association of the Lys420 polymorphism with physician-diagnosed asthma was of borderline statistical significance, and non-significant for Ser368. *SPINK5* codes for a 15 domain serine proteinase inhibitor, LEKT1, which inhibits plasmin, subtilisin A, cathepsin G, elastase, and trypsin (Mitsudo *et al.*, 2003). Polymorphisms in this inhibitor may affect the development of allergic disease by perturbing T and B cell maturation, mast cell activity, antigen handling or failure to inactivate the proteinase activity of allergens (Walley *et al.*, 2001).

Chronic obstructive pulmonary disease (COPD)

COPD is defined as airflow obstruction that does not vary significantly over time. It is a heterogeneous disease which may develop from destruction of the lung parenchyma (emphysema) or inflammation of the small airways (bronchiolitis). In any affected individual there may be considerable overlap, with each of these components contributing a varying amount to airflow obstruction. Chronic bronchitis is defined by the presence of cough with sputum production that is present for three or more months of the year, for at least two consecutive years. The most common cause of COPD is cigarette smoking (Anderson and Ferris, 1962; Fletcher and Peto, 1977; Surgeon General, 1979). Although chronic bronchitis is associated with cigarette smoking, and is often present in patients with COPD, it is not synonymous with COPD. The terms should not be used interchangeably as chronic bronchitis may occur in the absence of airflow obstruction.

Overlap between asthma and smoking-related COPD

The distinction between asthma and COPD is not always easy. Both conditions are associated with symptoms of wheeze and breathlessness and both are characterized by the presence of

airflow obstruction. Moreover, in long standing asthma there may be a loss of reversibility of airflow obstruction and smokers with bronchial hyper-reactivity are at greater risk of developing COPD. It has been suggested that the two conditions are part of the same disease spectrum, the so called ''Dutch hypothesis.'' There are however distinct differences in the pathophysiology of the two conditions both in the predominant type of inflammatory cell ($CD4^+$ lymphocytes and eosinophils in asthma and $CD8^+$ lymphocytes, macrophages and neutrophils in COPD) and the predominant site of inflammation (large and small airways in asthma and small airway and lung parenchyma in COPD) (Jeffery, 2001). Thus while the clinical features of asthma and COPD overlap, the differences in the pathophysiology suggests that they are unlikely to share a common genetic aetiology.

Environmental risk factors for the development of COPD: smoking

COPD is a major cause of global morbidity and mortality. An estimated 44 million people suffered from COPD in 1990 (Wise, 1997) and this condition accounts for approximately 2.8 million deaths/year. COPD is currently the fifth most common cause of death worldwide. However with the increase in the prevalence of smoking in the developing world (Peto *et al.*, 1999), it is predicted to become the third most common cause of death by 2020. The most important environmental risk factor for the development of COPD is cigarette smoking. Cross-sectional population studies have shown that smokers have poorer lung function (lower FEV_1) and a greater rate of decline in FEV_1 than non-smokers (Dockery *et al.*, 1988; Fletcher and Peto, 1977; Prescott *et al.*, 1997). However it remains unknown why only 15–20% of smokers develop significant airflow obstruction. One explanation is the variation in pack years smoked (one pack year is 20 cigarettes/day for one year), with heavy smokers being at significantly greater risk than light smokers. This is supported by population studies which have shown an association between cumulative pack years smoked and FEV_1 (Burrows *et al.*, 1977*a*) and rate of decline in FEV_1 (Jaakkola *et al.*, 1991; Peat *et al.*, 1990). However the correlation between FEV_1 and pack years smoked is relatively weak ($r = 0.3$) (Burrows *et al.*, 1977*a*) with only 10% of the variation in FEV_1 in smokers being explained by the amount smoked. Indeed the modest increase in rate of decline in FEV_1, from 20–30 mls a year in non-smokers to 30–45 mls in the majority of smokers, is unlikely to result in the development of significant airflow obstruction, even after many years of smoking (Dockery *et al.*, 1988).

COPD can also develop in individuals who enter adult life without attaining their maximal lung function as they are more vulnerable to the modestly increased rate of decline in FEV_1 observed in the majority of cigarette smokers (Dockery *et al.*, 1988). Childhood smokers in particular may have a potentially greater risk of developing COPD as they have reduced lung growth and a premature onset in decline of FEV_1, relative to their non-smoking peers (Gold *et al.*, 1996). Thus age of first starting to smoke may explain some of the variability in susceptibility to COPD.

An alternative explanation for the development of COPD in a small proportion of smokers is individual variation in susceptibility to cigarette smoke among smokers. This is supported by the identification of a subgroup of susceptible smokers in longitudinal studies who show a greatly accelerated decline in FEV_1 (Burrows *et al.*, 1987; Fletcher and Peto, 1977). The identification of a subset of susceptible smokers gives is a clear indication that additional environmental or genetic factors interact with cigarette smoke to result in the development of COPD.

Other environmental risk factors that may predispose smokers to COPD

In addition to cigarette smoking a number of environmental factors have been identified which

Table 26.4. Environmental risk factors associated with the development of COPD

Low birth weight
Childhood respiratory tract infections
Exposure to passive smoke in childhood
Latent viral infection
Ocupation exposures:
Cadmium worker
Coal miners
Gold miners and other underground tunnel workers
Indoor pollution:
Kerosene heaters
Nitrogen dioxide
Particulates from burning biomass fuels
Nutritional Factors:
Low dietary intake of dietary anti-oxidants
e.g. vitamins C, E and carotenoids
Oily fish (omega-3 fatty acids)
Flavenoids

impact on lung function (FEV_1) and may therefore predispose to the development of COPD. These are summarized in Table 26.4. The identification of low birth weight (Barker *et al.*, 1991), childhood respiratory tract infections (Burrows *et al.*, 1977*b*; Colley *et al.*, 1973; Kiernan *et al.*, 1976) and childhood passive smoke exposure (Lebowitz *et al.*, 1987; Wang *et al.*, 1994) as factors that may contribute to the pathogenesis of COPD suggests that the disease may begin in childhood. However, these observations are often derived from retrospective studies and the results need to be interpreted with caution (Phelan, 1984). In adults, increased expression of adenovirus protein in smokers with emphysema has provided evidence that latent viral infection may amplify the inflammatory response to cigarette smoke and contribute to the development of emphysema (Retamales *et al.*, 2001).

The association between environmental pollution and lung disease has been recognized since the London smogs of the 1950s, where the greatest excess in mortality resulted from an increase in deaths related to chronic bronchitis (Ministry of Health, 1954; Logan, 1953). Moreover, urban dwellers had poorer lung function and more respiratory symptoms than rural dwellers (Holland and Reid, 1965). Following the introduction of the clean air act in the 1970s, negative studies may have led to the perception that the adverse effects of pollution on respiratory health had been eliminated (Bouhuys *et al.*, 1978; Sunyer, 2001). However, several recent studies have again reported an association between air pollution and reduced lung function (Abbey *et al.*, 1998; Ackermann-Liebrich *et al.*, 1997; Chestnut *et al.*, 1991; Detels *et al.*, 1991; Peters *et al.*, 1999; Raizenne *et al.*, 1996; Stern *et al.*, 1994; Tashkin *et al.*, 1994). Although there is consistent evidence for a cross-sectional association between acute exacerbations of respiratory symptoms and environmental pollution, the role of pollution in the development of COPD is less clear, and requires further assessment (Sunyer, 2001).

Occupational exposure to particulates, gases and fumes have been identified as being associated with an increased risk of airflow obstruction particularly in miners and other underground workers (Cockcroft *et al.*, 1982; Lewis *et al.*, 1996; Love and Miller, 1982; Oxman *et al.*, 1993; Ulvestad *et al.*, 2000). Occupational exposure to cadmium fumes has also been identified as being a risk factor for emphysema (Davison *et al.*, 1988). However, these associations need to be interpreted with caution as workers in these industries are often smokers and may have other environmental risk factors for COPD. Moreover these occupational exposures are unlikely to contribute significantly to the burden of disease in regions or countries where these industries are not prevalent. In contrast, exposure to domestic pollution from cooking fumes could potentially have a much greater effect on the population because of the larger number of exposed individuals. Use of kerosene heaters (Behera *et al.*, 1994; Piitulainen *et al.*, 1998), indoor levels of nitrogen dioxide from cooking fuels (Fischer *et al.*, 1985) and indoor exposure to particulates from the burning of biomass fuels in the developing world (Dossing *et al.*, 1994; Pandey, 1984) have been implicated

as risk factors for impaired pulmonary function or respiratory disease. However, the effect of some of these exposures is small and may be reversible.

Early in the last century, studies in populations who were malnourished demonstrated that food deprivation resulted in an alteration in respiratory function and premature emphysema at autopsy (Sridhar, 1999). Cigarette smokers have a low dietary intake of fresh fruit (Bolton-Smith, 1993; Subar and Harlan, 1993), an important source of the dietary anti-oxidant vitamin C. Vitamin C is the most abundant anti-oxidant in the extra-cellular fluid lining the lungs (Slade *et al.*, 1993). Cigarette tar contains a very high concentration of free radicals (Pryor and Stone, 1993), with each inhalation from a cigarette containing approximately 10^{16} oxidant particles (Church and Pryor, 1985). An imbalance between oxidative stress and anti-oxidant activity has been suggested as a possible mechanism for the development of COPD in smokers (Cross *et al.*, 1999; Macnee and Rahman, 1999; Traber *et al.*, 2000). Proposed cellular mechanisms by which this oxidant stress may lead to lung damage include direct damage to the lung tissues (elastin and collagen), increased phagocytic oxidant activity, activation of destructive enzymes (e.g. metalloproteinases) and the inactivation of natural inhibitors (e.g. α_1-antitrypsin). The role of nutritional factors is supported by cross-sectional population studies that have shown an association between respiratory disease or lung function and dietary anti-oxidant status as determined by self-reported diet (Britton *et al.*, 1995; Butland *et al.*, 2000; Chen *et al.*, 2001; Dow *et al.*, 1996; Grievink *et al.*, 1998; Hu *et al.*, 1998; Hu and Cassano, 2000; Schwartz and Weiss, 1990; Strachan *et al.*, 1991) or biomarkers of anti-oxidant activity (Hu *et al.*, 1998; Hu and Cassano, 2000; Ness *et al.*, 1996; Schwartz and Weiss, 1990). However the strength of the association demonstrated in these studies has been relatively weak, and only one cross-sectional study has reported an interaction between low vitamin C levels and the risk of airflow obstruction (Sargeant *et al.*, 2000). Anti-oxidant nutrients other than vitamin C and dietary fish oil have also been associated with lung function (Schunemann *et al.*, 2001; Shahar *et al.*, 1994; Tabak *et al.*, 2001). Although there is little prospective data for the role of diet in the development of COPD, ecological data suggests that differences in COPD mortality rates between countries may be partially explained by dietary differences in fish and fresh fruit consumption (Tabak *et al.*, 1998).

Genetic risk factors for the development of COPD

Alpha$_1$-antitrypsin deficiency

The observation that only a minority of cigarette smokers develop COPD suggests that additional factors contribute to the impact of smoking on the development of chronic airflow obstruction. The most important genetic factor in the development of COPD is the Z allele of α_1-antitrypsin. Alpha$_1$-antitrypsin is the most abundant circulating proteinase inhibitor and plays a central role in protecting the lungs against proteolytic attack. The Z allele (*Glu342Lys*) results in 85% of the synthesized mutant α_1-antitrypsin being retained as polymers within hepatocytes (Lomas *et al.*, 1992). This results in plasma levels of α_1-antitrypsin that are only 10–15% of the normal M allele. Homozygotes for the Z allele are greatly predisposed to developing emphysema if they smoke (Piitulainen and Eriksson, 1999). However severe Z α_1-antitrypsin deficiency makes up only 1–2% of all cases of COPD and there is considerable variability in FEV_1 between current and ex-smokers with the same PI Z genotype (Silverman *et al.*, 1990). This suggests that other co-existing genetic factors must predispose to lung disease in PI Z individuals.

A logical follow-on from the association of α_1-antitrypsin deficiency with emphysema is an assessment of the risk of COPD in heterozygotes who carry an abnormal Z allele and a normal M allele. These individuals have plasma α_1-antitrypsin levels that are approximately 65% of normal. A population-based study demonstrated that PI MZ heterozygotes do not have a clearly increased

risk of lung damage (Bruce *et al.*, 1984). However if groups of patients are collected who already have COPD, then the prevalence of PI MZ individuals appears to be elevated (Lieberman *et al.*, 1986). In addition, a longitudinal study has demonstrated that among COPD patients, the PI MZ heterozygotes have a more rapid decline in lung function (Tarjan *et al.*, 1994). These data suggest that either all PI MZ individuals are at slightly increased risk for the development of COPD, or that a subset of the PI MZ subjects are at substantially increased risk of pulmonary damage if they smoke. An alternative explanation is that the apparent increased risk among PI MZ subjects reflects ascertainment bias, and the elevated rate of PI MZ subjects among COPD cases reflects the influence of other, as yet unidentified, factors.

Other genetic risk factors for the development of COPD

The identification of α_1-antitrypsin deficiency as a risk factor for the development of emphysema provided clear evidence of a genetic component for the development of COPD. However, α_1-antitrypsin deficiency is found in only 1–2% of all cases of COPD. Two recent studies have provided evidence to suggest that genetic factors other than α_1-antitrypsin deficiency may be important in the development of COPD. Both studies assessed the familial risk of airflow obstruction in smoking and ex-smoking relatives of index cases (probands) who had early onset COPD without severe α_1-antitrypsin deficiency. Silverman *et al.* found current and ex-smoking first degree relatives of probands had a three-fold greater risk of a reduced FEV_1 relative to current or ex-smoking controls (Silverman *et al.*, 1998). More recently McCloskey *et al.* assessed the risk of airflow obstruction in 126 current and ex-smoking siblings of probands relative to 419 matched population based controls (McCloskey *et al.*, 2001). The prevalence of airflow obstruction was far greater in the sibling group (31.5%) than the control group (9.3%), with the siblings having a 4.7-fold greater risk of airflow obstruction. This was despite the two groups being very similar in respect of age, sex and, most importantly, smoking history.

These two studies provide strong evidence for genetic risk factors other than α_1-antitrypsin deficiency in the development of COPD. The use of population-based controls in the second study gave an estimate for λ_s of 4.7, for the development of smoking-related COPD. In both studies the probands had severe early onset disease which reduced the likelihood of phenocopies. Moreover, the emphysema phenotype was confirmed by computer tomography imaging of the lungs in the majority of probands. Both studies found the non-smoking siblings of the probands to have normal lung function, emphasizing the importance of the interaction between genetic predisposition and exposure to cigarette smoke.

A number of association studies have assessed polymorphisms in candidate genes that may predispose to the development of COPD. In these studies, the frequency of an allele in affected individuals is compared with the frequency in a control (unaffected) population. Table 26.5 summarizes the genes for which there is the strongest evidence of a link with COPD. In addition to anti-proteinases, proteins that are important in detoxifying cigarette smoke or modulating the inflammatory response have also been studied. Xenobiotic metabolizing enzymes, for example, may play an important role in metabolizing reactive substances (microsomal epoxide hydrolase) or aromatic hydrocarbons (glutathione S-transferase) found in cigarette smoke. A high number of dinucleotide repeats (>30) in the 5′-flanking region of the heme-oxygenase-1 (*HMOX-1*) gene has been reported to be associated with COPD. *HMOX-1* protects against cellular oxidant injury by degrading heme to biliverdin. The association of this polymorphism with COPD suggests that this enzyme may also be important in protecting against non-heme mediated oxidant injury (Yamada *et al.*, 2000). TNF-α is a pro-inflammatory cytokine. The TNF-2 polymorphism has been associated with increased levels of TNF-α

Table 26.5. Examples of candidate genes that have been associated with COPD

Gene	Association
Anti-proteinases	
Alpha-1-antitrypsin	Greater frequency of g→a polymorphism in the 3′ non-coding region reported in subjects with COPD (Kalsheker *et al.*, 1987; Poller *et al.*, 1990)
Alpha-1-antichymotrypsin	Leu^{55}→Pro mutation (Poller *et al.*, 1993) and a Pro^{229}→Ala (Poller *et al.*, 1992) mutation have been reported to be more common in individuals with COPD than controls
Xenobiotic metabolizing enzymes	
Microsomal epoxide hydrolase (*mEPHX*)	The slow metabolizing polymorphism of *mEPHX* has been reported to be more common in people with COPD (Smith and Harrison, 1997), and to be associated with more severe COPD (Yoshikawa *et al.*, 2000).
Glutathione S-transferase (*GST*)	The Null allele of *GST M1* (*GSTM1 0/0*) is reported to be associated with chronic bronchitis in smokers (Baranova *et al.*, 1997), and with the presence of emphysema in patients with lung cancer (Harrison *et al.*, 1997). Heterozygotes for the *GSTP1/1le 105* polymorphism of the *GSTP1* gene have been reported to be more prevalent in patients with COPD (Ishii *et al.*, 1999).
Antioxidants	
Heme-oxygenase-1 (*HMOX-1*)	A microsatellite polymorphism in the 5′ region of the *HMOX-1* gene was reported to be more common in patients with emphysema than controls (Yamada *et al.*, 2000).
Inflammatory cytokines	
Tumour necrosis factor alpha (TNF-α)	TNF-2 polymorphism reported to be more common in patients with airflow obstruction (Sakao *et al.*, 2001).

in vitro, and theoretically therefore could be associated with a greater inflammatory response in the lungs of smokers. Although there is a plausible biological mechanism for each of these genes being involved in the pathogenesis of COPD, the evidence is not conclusive for any of those considered. Moreover other studies have failed to reproduce the associations (Benetazzo *et al.*, 1999; Sandford *et al.*, 1997; 1998; Yim *et al.*, 2000).

Association studies in individuals with COPD are often limited by poor phenotype definition, small sample size and, in particular, poor selection of controls. To avoid spurious results controls should be matched to cases as closely as possible by age, sex, smoking history, ethnic background and occupation. However this is rarely the case and therefore the results of these studies need to be interpreted with caution. In addition to these methodological concerns there are two other limitations of association studies. Firstly, they only allow the study of genes that have been previously described and will not therefore allow

the identification of novel genes that may predispose to COPD. Secondly, they require the study population to be dichotomized into affected and unaffected phenotypes. However in contrast to asthma, spirometric indices such as FEV_1 and FEV_1/FVC ratio provide useful intermediate phenotypes for quantitative trait analysis in studying the genetics of COPD.

These limitations have resulted in alternative methods being employed to study the genetics of COPD. Using linkage analysis, a recent study in families with early onset COPD identified significant linkage between the FEV_1/FVC ratio and chromosome 2q, and suggested linkage between FEV_1 and chromosome 12p (Silverman *et al.*, 2002). A large international study now has collected over 700 hundred affected sib pairs from around the world for linkage analysis. In addition to spirometric data, the individuals in this study have undergone CT scans of the chest to allow quantitative assessment of emphysema, thus allowing precise identification of the phenotype. The first round of genome scans are due to be performed in the near future and it is hoped that this will identify novel genes that predispose to the development of airways disease and/or emphysema.

Conclusions

Asthma and chronic obstructive pulmonary disease (COPD) are common causes of airflow obstruction. There is strong evidence to suggest a genetic basis for both conditions, yet to date α_1-antitrypsin deficiency remains the only proven genetic risk factor for the development of airflow obstruction. The identification of other genes is complicated by multiple phenotypes, polygenicity, incomplete penetrance and a high phenocopy rate. Moreover, the role of cigarette smoking in COPD and allergen exposure in asthma clearly demonstrate the importance of gene–environment interactions in both conditions. The importance of accurate identification of the phenotype and the measurement of confounding environmental exposures in future studies cannot be overstated. Thus progress in understanding the genetic basis of these diseases will require the integration of traditional epidemiological methods with modern genetic research methodology.

REFERENCES

Abbey, D. E., Burchette, R. J., Knutsen, S. F. *et al.* (1998). Long-term particulate and other air pollutants and lung function in nonsmokers. *Am J Respir Crit Care Med*, **158**, 289–98.

Ackermann-Liebrich, U., Leuenberger, P., Schwartz, J. *et al.* (1997). Lung function and long term exposure to air pollutants in Switzerland. Study on Air Pollution and Lung Diseases in Adults (SAPALDIA) Team. *Am J Respir Crit Care Med*, **155**, 122–9.

Anderson, D. O. and Ferris, B. G., Jr. (1962). Role of tobacco smoking in the causation of chronic respiratory disease. *N Engl J Med*, **267**, 787–94.

Anderson, H. R., Butland, B. K. and Strachan, D. P. (1994). Trends in prevalence and severity of childhood asthma. *Br Med J*, **308**, 1600–4.

Baranova, H., Perriot, J., Albuisson, E. *et al.* (1997). Peculiarities of the GSTM1 0/0 genotype in French heavy smokers with various types of chronic bronchitis. *Hum Genet*, **99**, 822–6.

Barker, D. J., Godfrey, K. M., Fall, C. *et al.* (1991). Relation of birth weight and childhood respiratory infection to adult lung function and death from chronic obstructive airways disease. *Bmj*, **303**, 671–5.

Behera, D., Jindal, S. K. and Malhotra, H. S. (1994). Ventilatory function in nonsmoking rural Indian women using different cooking fuels. *Respiration*, **61**, 89–92.

Benetazzo, M. G., Gile, L. S., Bombieri, C. *et al.* (1999). alpha 1-antitrypsin TAQ I polymorphism and alpha 1-antichymotrypsin mutations in patients with obstructive pulmonary disease. *Respir Med*, **93**, 648–54.

Bodner, C., Godden, D., Brown, K. *et al.* (1999). Antioxidant intake and adult-onset wheeze: a case-control study. Aberdeen WHEASE Study Group. *Eur Respir J*, **13**, 22–30.

Bodner, C., Godden, D. and Seaton, A. (1998). Family size, childhood infections and atopic diseases. The Aberdeen WHEASE Group. *Thorax*, **53**, 28–32.

Bolton-Smith, C. (1993). Antioxidant vitamin intakes in Scottish smokers and nonsmokers. Dose effects and

biochemical correlates. *Ann N Y Acad Sci*, **686**, 347–58; discussion 358–60.

Bouhuys, A., Beck, G. J. and Schoenberg, J. B. (1978). Do present levels of air pollution outdoors affect respiratory health? *Nature*, **276**, 466–71.

British Thoracic Society (2003). The burden of lung disease. A statistics report from the British Thoracic Society, The British Thoracic Society, London.

British Thoracic Society; Scottish Intercollegiate Guidelines Network. (2003). British guideline on the management of asthma. *Thorax*, **58**, Suppl 1, i1–94.

Britton, J., Pavord, I., Richards, K. *et al.* (1994). Dietary magnesium, lung function, wheezing, and airway hyperreactivity in a random adult population sample. *Lancet*, **344**, 357–62.

Britton, J. R., Pavord, I. D., Richards, K. A. *et al.* (1995). Dietary antioxidant vitamin intake and lung function in the general population. *Am J Respir Crit Care Med*, **151**, 1383–7.

Bruce, R. M., Cohen, B. H., Diamond, E. L. *et al.* (1984). Collaborative study to assess risk of lung disease in Pi MZ phenotype subjects. *Am Rev Respir Dis*, **130**, 386–90.

Burney, P. G., Chinn, S. and Rona, R. J. (1990). Has the prevalence of asthma increased in children? Evidence from the national study of health and growth 1973–86. *Bmj*, **300**, 1306–10.

Burney, P. G. J. (1992). *Asthma*. London: Chapman & Hall.

Burrows, B., Knudson, R. J., Camilli, A. E., Lyle, S. K. and Lebowitz, M. D. (1987). The ''horse-racing effect'' and predicting decline in forced expiratory volume in one second from screening spirometry. *Am Rev Respir Dis*, **135**, 788–93.

Burrows, B., Knudson, R. J., Cline, M. G. and Lebowitz, M. D. (1977*a*). Quantitative relationships between cigarette smoking and ventilatory function. *Am Rev Respir Dis*, **115**, 195–205.

Burrows, B., Knudson, R. J. and Lebowitz, M. D. (1977*b*). The relationship of childhood respiratory illness to adult obstructive airway disease. *Am Rev Respir Dis*, **115**, 751–60.

Butland, B. K., Fehily, A. M. and Elwood, P. C. (2000). Diet, lung function, and lung function decline in a cohort of 2512 middle aged men. *Thorax*, **55**, 102–8.

Chavanas, S., Bodemer, C., Rochat, A. *et al.* (2000). Mutations in *SPINK5*, encoding a serine protease inhibitor, cause Netherton syndrome. *Nat Genet*, **25**, 141–2.

Chen, R., Tunstall-Pedoe, H., Bolton-Smith, C., Hannah, M. K. and Morrison, C. (2001). Association of dietary antioxidants and waist circumference with pulmonary function and airway obstruction. *Am J Epidemiol*, **153**, 157–63.

Chestnut, L. G., Schwartz, J., Savitz, D. A. and Burchfiel, C. M. (1991). Pulmonary function and ambient particulate matter: epidemiological evidence from NHANES I. *Arch Environ Health*, **46**, 135–44.

Church, D. F. and Pryor, W. A. (1985). Free-radical chemistry of cigarette smoke and its toxicological implications. *Environ Health Perspect*, **64**, 111–26.

Cockcroft, A., Seal, R. M., Wagner, J. C. *et al.* (1982). Postmortem study of emphysema in coalworkers and noncoalworkers. *Lancet*, **2**, 600–3.

Colley, J. R., Douglas, J. W. and Reid, D. D. (1973). Respiratory disease in young adults: influence of early childhood lower respiratory tract illness, social class, air pollution, and smoking. *Br Med J*, **3**, 195–8.

Cross, C. E., Traber, M., Eiserich, J. and van der Vliet, A. (1999). Micronutrient antioxidants and smoking. *Br Med Bull*, **55**, 691–704.

Davison, A. G., Fayers, P. M., Taylor, A. J. *et al.* (1988). Cadmium fume inhalation and emphysema. *Lancet*, **1**, 663–7.

Department of Health and Social Services (2003). Report on diagnosis and treatment of asthma DHSS 1992, Published by Department of Health and Social Services (DHSS), London.

Detels, R., Tashkin, D. P., Sayre, J. W. *et al.* (1991). The UCLA population studies of CORD: X. A cohort study of changes in respiratory function associated with chronic exposure to SOx, NOx, and hydrocarbons. *Am J Public Health*, **81**, 350–9.

Dockery, D. W., Speizer, F. E., Ferris, B. G., Jr. *et al.* (1988). Cumulative and reversible effects of lifetime smoking on simple tests of lung function in adults. *Am Rev Respir Dis*, **137**, 286–92.

Dossing, M., Khan, J. and al-Rabiah, F. (1994). Risk factors for chronic obstructive lung disease in Saudi Arabia. *Respir Med*, **88**, 519–22.

Dow, L., Tracey, M., Villar, A. *et al.* (1996). Does dietary intake of vitamins C and E influence lung function in older people? *Am J Respir Crit Care Med*, **154**, 1401–4.

Edfors-Lubs, M. L. (1971). Allergy in 7000 twin pairs. *Acta Allergol*, **26**, 249–85.

Fischer, P., Remijn, B., Brunekreef, B. *et al.* (1985). Indoor air pollution and its effect on pulmonary function of adult non-smoking women: II. Associations between nitrogen dioxide and pulmonary function. *Int J Epidemiol*, **14**, 221–6.

Flatt, A., Pearce, N., Thomson, C. D. *et al.* (1990). Reduced selenium in asthmatic subjects in New Zealand. *Thorax*, **45**, 95–9.

Fletcher, C. and Peto, R. (1977). The natural history of chronic airflow obstruction. *Br Med J*, **1**, 1645–8.

Fogarty, A. and Britton, J. (2000). The role of diet in the aetiology of asthma. *Clin Exp Allergy*, **30**, 615–27.

Forastiere, F., Pistelli, R., Sestini, P. *et al.* (2000). Consumption of fresh fruit rich in vitamin C and wheezing symptoms in children. SIDRIA Collaborative Group, Italy (Italian Studies on Respiratory Disorders in Children and the Environment). *Thorax*, **55**, 283–8.

Gold, D. R., Wang, X., Wypij, D. *et al.* (1996). Effects of cigarette smoking on lung function in adolescent boys and girls. *N Engl J Med*, **335**, 931–7.

Grievink, L., Smit, H. A., Ocke, M. C., van't Veer, P. and Kromhout, D. (1998). Dietary intake of antioxidant (pro)-vitamins, respiratory symptoms and pulmonary function: the MORGEN study. *Thorax*, **53**, 166–71.

Grievink, L., Zijlstra, A. G., Ke, X. and Brunekreef, B. (1999). Double-blind intervention trial on modulation of ozone effects on pulmonary function by antioxidant supplements. *Am J Epidemiol*, **149**, 306–14.

Harrison, D. J., Cantlay, A. M., Rae, F., Lamb, D. and Smith, C. A. (1997). Frequency of glutathione S-transferase M1 deletion in smokers with emphysema and lung cancer. *Hum Exp Toxicol*, **16**, 356–60.

Heinzmann, A. and Deichmann, K. A. (2001). Genes for atopy and asthma. *Curr Opin Allergy Clin Immunol*, **1**, 387–92.

Hijazi, N., Abalkhail, B. and Seaton, A. (2000). Diet and childhood asthma in a society in transition: a study in urban and rural Saudi Arabia. *Thorax*, **55**, 775–9.

Hodge, L., Salome, C. M., Peat, J. K. *et al.* (1996). Consumption of oily fish and childhood asthma risk. *Med J Aust*, **164**, 137–40.

Holland, W. W. and Reid, D. D. (1965). The urban factor in chronic bronchitis. *Lancet*, **40**, 445–8.

Hu, G. and Cassano, P. A. (2000). Antioxidant nutrients and pulmonary function: the Third National Health and Nutrition Examination Survey (NHANES III). *Am J Epidemiol*, **151**, 975–81.

Hu, G., Zhang, X., Chen, J. *et al.* (1998). Dietary vitamin C intake and lung function in rural China. *Am J Epidemiol*, **148**, 594–9.

Ishii, T., Matsuse, T., Teramoto, S. *et al.* (1999). Glutathione S-transferase P1 (GSTP1) polymorphism in patients with chronic obstructive pulmonary disease. *Thorax*, **54**, 693–6.

Jaakkola, M. S., Ernst, P., Jaakkola, J. J. *et al.* (1991). Effect of cigarette smoking on evolution of ventilatory lung function in young adults: an eight year longitudinal study. *Thorax*, **46**, 907–13.

Jarvis, D., Chinn, S., Luczynska, C. and Burney, P. (1997). The association of family size with atopy and atopic disease. *Clin Exp Allergy*, **27**, 240–5.

Jeffery, P. K. (2001). Remodeling in asthma and chronic obstructive lung disease. *Am J Respir Crit Care Med*, **164**, S28–38.

Kalsheker, N. A., Hodgson, I. J., Watkins, G. L. *et al.* (1987). Deoxyribonucleic acid (DNA) polymorphism of the alpha 1-antitrypsin gene in chronic lung disease. *Br Med J (Clin Res Ed)*, **294**, 1511–14.

Keeley, D. J., Neill, P. and Gallivan, S. (1991). Comparison of the prevalence of reversible airways obstruction in rural and urban Zimbabwean children. *Thorax*, **46**, 549–53.

Kiernan, K. E., Colley, J. R., Douglas, J. W. and Reid, D. D. (1976). Chronic cough in young adults in relation to smoking habits, childhood environment and chest illness. *Respiration*, **33**, 236–44.

Lebowitz, M. D., Holberg, C. J., Knudson, R. J. and Burrows, B. (1987). Longitudinal study of pulmonary function development in childhood, adolescence, and early adulthood. Development of pulmonary function. *Am Rev Respir Dis*, **136**, 69–75.

Lewis, S., Bennett, J., Richards, K. and Britton, J. (1996). A cross sectional study of the independent effect of occupation on lung function in British coal miners. *Occup Environ Med*, **53**, 125–8.

Lieberman, J., Winter, B. and Sastre, A. (1986). Alpha 1-antitrypsin Pi-types in 965 COPD patients. *Chest*, **89**, 370–3.

Logan, W. P. (1953). Mortality in the London fog incident, 1952. *Lancet*, **1**, 336–8.

Lomas, D. A., Evans, D. L., Finch, J. T. and Carrell, R. W. (1992). The mechanism of Z alpha 1-antitrypsin accumulation in the liver. *Nature*, **357**, 605–7.

Love, R. G. and Miller, B. G. (1982). Longitudinal study of lung function in coal-miners. *Thorax*, **37**, 193–7.

Macnee, W. and Rahman, I. (1999). Oxidants and antioxidants as therapeutic targets in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*, **160**, S58–65.

Matricardi, P. M., Franzinelli, F., Franco, A. *et al.* (1998). Sibship size, birth order, and atopy in 11,371 Italian young men. *J Allergy Clin Immunol*, **101**, 439–44.

Mattes, J., Karmaus, W., Moseler, M., Frischer, T. and Kuehr, J. (1998). Accumulation of atopic disorders

within families: a sibling effect only in the offspring of atopic fathers. *Clin Exp Allergy*, **28**, 1480–6.

McCloskey, S.C., Patel, B.D., Hinchliffe, S.J. *et al.* (2001). Siblings of patients with severe chronic obstructive pulmonary disease have a significant risk of airflow obstruction. *Am J Respir Crit Care Med*, **164**, 1419–24.

Miedema, I., Feskens, E.J., Heederik, D. and Kromhout, D. (1993). Dietary determinants of long-term incidence of chronic nonspecific lung diseases. The Zutphen Study. *Am J Epidemiol*, **138**, 37–45.

Ministry of Health (1954). Mortality and morbidity during the London fog of December 1952. London: HMSO.

Mitsudo, K., Jayakumar, A., Henderson, Y. *et al.* (2003). Inhibition of serine proteinases plasmin, trypsin, subtilisin A, cathepsin G, and elastase by LEKTI: a kinetic analysis. *Biochemistry*, **42**, 3874–81.

National Heart, Lung, and Blood Institute (1992). International consensus report on diagnosis and treatment of asthma. National Heart, Lung, and Blood Institute, National Institutes of Health. Bethesda, Maryland 20892. Publication no. 92–3091, March 1992. *Eur Respir J*, **5**, 601–41.

Ness, A.R., Khaw, K.T., Bingham, S. and Day, N.E. (1996). Vitamin C status and respiratory function. *Eur J Clin Nutr*, **50**, 573–9.

Oxman, A.D., Muir, D.C., Shannon, H.S. *et al.* (1993). Occupational dust exposure and chronic obstructive pulmonary disease. A systematic overview of the evidence. *Am Rev Respir Dis*, **148**, 38–48.

Palmer, L.J. (2001). Linkages and associations to intermediate phenotypes underlying asthma and allergic disease. *Curr Opin Allergy Clin Immunol*, **1**, 393–8.

Pandey, M.R. (1984). Domestic smoke pollution and chronic bronchitis in a rural community of the Hill Region of Nepal. *Thorax*, **39**, 337–9.

Peat, J.K., van den Berg, R.H., Green, W.F. *et al.* (1994). Changing prevalence of asthma in Australian children. *Bmj*, **308**, 1591–6.

Peat, J.K., Woolcock, A.J. and Cullen, K. (1990). Decline of lung function and development of chronic airflow limitation: a longitudinal study of non-smokers and smokers in Busselton, Western Australia. *Thorax*, **45**, 32–7.

Peters, J.M., Avol, E., Gauderman, W.J. *et al.* (1999). A study of twelve Southern California communities with differing levels and types of air pollution. II. Effects on pulmonary function. *Am J Respir Crit Care Med*, **159**, 768–75.

Peto, R., Chen, Z.M. and Boreham, J. (1999). Tobacco – the growing epidemic. *Nat Med*, **5**, 15–17.

Phelan, P.D. (1984). Does adult chronic obstructive lung disease really begin in childhood? *Br J Dis Chest*, **78**, 1–9.

Piitulainen, E. and Eriksson, S. (1999). Decline in FEV1 related to smoking status in individuals with severe alpha1-antitrypsin deficiency (PiZZ). *Eur Respir J*, **13**, 247–51.

Piitulainen, E., Tornling, G. and Eriksson, S. (1998). Environmental correlates of impaired lung function in non-smokers with severe alpha 1-antitrypsin deficiency (PiZZ). *Thorax*, **53**, 939–43.

Poller, W., Faber, J.P., Scholz, S. *et al.* (1992). Mis-sense mutation of alpha 1-antichymotrypsin gene associated with chronic lung disease. *Lancet*, **339**, 1538.

Poller, W., Faber, J.P., Weidinger, S. *et al.* (1993). A leucine-to-proline substitution causes a defective alpha 1-antichymotrypsin allele associated with familial obstructive lung disease. *Genomics*, **17**, 740–3.

Poller, W., Meisen, C. and Olek, K. (1990). DNA polymorphisms of the alpha 1-antitrypsin gene region in patients with chronic obstructive pulmonary disease. *Eur J Clin Invest*, **20**, 1–7.

Prescott, E., Bjerg, A.M., Andersen, P.K., Lange, P. and Vestbo, J. (1997). Gender difference in smoking effects on lung function and risk of hospitalization for COPD: results from a Danish longitudinal population study. *Eur Respir J*, **10**, 822–7.

Pryor, W.A. and Stone, K. (1993). Oxidants in cigarette smoke. Radicals, hydrogen peroxide, peroxynitrate, and peroxynitrite. *Ann N Y Acad Sci*, **686**, 12–27; discussion 27–8.

Raizenne, M., Neas, L.M., Damokosh, A.I. *et al.* (1996). Health effects of acid aerosols on North American children: pulmonary function. *Environ Health Perspect*, **104**, 506–14.

Rautalahti, M., Virtamo, J., Haukka, J. *et al.* (1997). The effect of alpha-tocopherol and beta-carotene supplementation on COPD symptoms. *Am J Respir Crit Care Med*, **156**, 1447–52.

Retamales, I., Elliott, W.M., Meshi, B. *et al.* (2001). Amplification of inflammation in emphysema and its association with latent adenoviral infection. *Am J Respir Crit Care Med*, **164**, 469–73.

Roemer, W., Hoek, G. and Brunekreef, B. (1993). Effect of ambient winter air pollution on respiratory health of children with chronic respiratory symptoms. *Am Rev Respir Dis*, **147**, 118–24.

Romagnani, S. (1992). Human TH1 and TH2 subsets: regulation of differentiation and role in protection and immunopathology. *Int Arch Allergy Immunol*, **98**, 279–85.

Romieu, I., Meneses, F., Ramirez, M. *et al.* (1998). Antioxidant supplementation and respiratory functions among workers exposed to high levels of ozone. *Am J Respir Crit Care Med*, **158**, 226–32.

Rona, R. J., Duran-Tauleria, E. and Chinn, S. (1997). Family size, atopic disorders in parents, asthma in children, and ethnicity. *J Allergy Clin Immunol*, **99**, 454–60.

Sakao, S., Tatsumi, K., Igari, H. *et al.* (2001). Association of tumor necrosis factor alpha gene promoter polymorphism with the presence of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*, **163**, 420–2.

Sandford, A. J., Chagani, T., Weir, T. D. and Pare, P. D. (1998). Alpha 1-antichymotrypsin mutations in patients with chronic obstructive pulmonary disease. *Dis Markers*, **13**, 257–60.

Sandford, A. J., Spinelli, J. J., Weir, T. D. and Pare, P. D. (1997). Mutation in the 3′ region of the alpha-1-antitrypsin gene and chronic obstructive pulmonary disease. *J Med Genet*, **34**, 874–5.

Sargeant, L. A., Jaeckel, A. and Wareham, N. J. (2000). Interaction of vitamin C with the relation between smoking and obstructive airways disease in EPIC Norfolk. European Prospective Investigation into Cancer and Nutrition. *Eur Respir J*, **16**, 397–403.

Schunemann, H. J., Grant, B. J., Freudenheim, J. L. *et al.* (2001). The relation of serum levels of antioxidant vitamins C and E, retinol and carotenoids with pulmonary function in the general population. *Am J Respir Crit Care Med*, **163**, 1246–55.

Schwartz, J. and Weiss, S. T. (1990). Dietary factors and their relation to respiratory symptoms. The Second National Health and Nutrition Examination Survey. *Am J Epidemiol*, **132**, 67–76.

Seaton, A., Godden, D. J. and Brown, K. (1994). Increase in asthma: a more toxic environment or a more susceptible population? *Thorax*, **49**, 171–4.

Shahar, E., Folsom, A. R., Melnick, S. L. *et al.* (1994). Dietary n-3 polyunsaturated fatty acids and smoking-related chronic obstructive pulmonary disease. Atherosclerosis Risk in Communities Study Investigators. *N Engl J Med*, **331**, 228–33.

Shapiro, S. D. and Owen, C. A. (2002). *ADAM-33* surfaces as an asthma gene. *N Engl J Med*, **347**, 936–8.

Silverman, E. K., Chapman, H. A., Drazen, J. M. *et al.* (1998). Genetic epidemiology of severe, early-onset chronic obstructive pulmonary disease. Risk to relatives for airflow obstruction and chronic bronchitis. *Am J Respir Crit Care Med*, **157**, 1770–8.

Silverman, E. K., Palmer, L. J., Mosley, J. D. *et al.* (2002). Genomewide linkage analysis of quantitative spirometric phenotypes in severe early-onset chronic obstructive pulmonary disease. *Am J Hum Genet*, **70**, 1229–39.

Silverman, E. K., Province, M. A., Campbell, E. J., Pierce, J. A. and Rao, D. C. (1990). Biochemical intermediates in alpha 1-antitrypsin deficiency: residual family resemblance for total alpha 1-antitrypsin, oxidized alpha 1-antitrypsin, and immunoglobulin E after adjustment for the effect of the Pi locus. *Genet Epidemiol*, **7**, 137–49.

Slade, R., Crissman, K., Norwood, J. and Hatch, G. (1993). Comparison of antioxidant substances in bronchoalveolar lavage cells and fluid from humans, guinea pigs, and rats. *Exp Lung Res*, **19**, 469–84.

Smith, C. A. and Harrison, D. J. (1997). Association between polymorphism in gene for microsomal epoxide hydrolase and susceptibility to emphysema. *Lancet*, **350**, 630–3.

Soutar, A., Seaton, A. and Brown, K. (1997). Bronchial reactivity and dietary antioxidants. *Thorax*, **52**, 166–70.

Sridhar, M. K. (1999). Nutrition and lung health. *Proc Nutr Soc*, **58**, 303–8.

Stern, B. R., Raizenne, M. E., Burnett, R. T. *et al.* (1994). Air pollution and childhood respiratory health: exposure to sulfate and ozone in 10 Canadian rural communities. *Environ Res*, **66**, 125–42.

Stone, J., Hinks, L. J., Beasley, R., Holgate, S. T. and Clayton, B. A. (1989). Reduced selenium status of patients with asthma. *Clin Sci (Lond)*, **77**, 495–500.

Strachan, D. P. (1989). Hay fever, hygiene, and household size. *Bmj*, **299**, 1259–60.

Strachan, D. P., Cox, B. D., Erzinclioglu, S. W., Walters, D. E. and Whichelow, M. J. (1991). Ventilatory function and winter fresh fruit consumption in a random sample of British adults. *Thorax*, **46**, 624–9.

Strachan, D. P., Harkins, L. S. and Golding, J. (1997*a*). Sibship size and self-reported inhalant allergy among adult women. ALSPAC Study Team. *Clin Exp Allergy*, **27**, 151–5.

Strachan, D. P., Harkins, L. S., Johnston, I. D. and Anderson, H. R. (1997*b*). Childhood antecedents of

allergic sensitization in young British adults. *J Allergy Clin Immunol*, **99**, 6–12.

Subar, A. F. and Harlan, L. C. (1993). Nutrient and food group intake by tobacco use status: the 1987 National Health Interview Survey. *Ann N Y Acad Sci*, **686**, 310–21; discussion 321–2.

Sunyer, J. (2001). Urban air pollution and chronic obstructive pulmonary disease: a review. *Eur Respir J*, **17**, 1024–33.

Surgeon General. (1979). Smoking and Health. A report of the Surgeon General. DHEW publication no (PHS) 79–50066. Washington, DC: US GPO.

Svanes, C., Jarvis, D., Chinn, S. and Burney, P. (1999). Childhood environment and adult atopy: results from the European Community Respiratory Health Survey. *J Allergy Clin Immunol*, **103**, 415–20.

Tabak, C., Arts, I. C., Smit, H. A., Heederik, D. and Kromhout, D. (2001). Chronic obstructive pulmonary disease and intake of catechins, flavonols, and flavones: the MORGEN Study. *Am J Respir Crit Care Med*, **164**, 61–4.

Tabak, C., Feskens, E. J., Heederik, D. *et al.* (1998). Fruit and fish consumption: a possible explanation for population differences in COPD mortality (The Seven Countries Study). *Eur J Clin Nutr*, **52**, 819–25.

Tarjan, E., Magyar, P., Vaczi, Z., Lantos, A. and Vaszar, L. (1994). Longitudinal lung function study in heterozygous PiMZ phenotype subjects. *Eur Respir J*, **7**, 2199–204.

Tashkin, D. P., Detels, R., Simmons, M. *et al.* (1994). The UCLA population studies of chronic obstructive respiratory disease: XI. Impact of air pollution and smoking on annual change in forced expiratory volume in one second. *Am J Respir Crit Care Med*, **149**, 1209–17.

The European Community Respiratory Health Survey Group (1997). Genes for asthma? An analysis of the European Community Respiratory Health Survey. *Am J Respir Crit Care Med*, **156**, 1773–80.

Traber, M. G., van der Vliet, A., Reznick, A. Z. and Cross, C. E. (2000). Tobacco-related diseases. Is there a role for antioxidant micronutrient supplementation? *Clin Chest Med*, **21**, 173–87.

Troisi, R. J., Willett, W. C., Weiss, S. T. *et al.* (1995). A prospective study of diet and adult-onset asthma. *Am J Respir Crit Care Med*, **151**, 1401–8.

Ulvestad, B., Bakke, B., Melbostad, E. *et al.* (2000). Increased risk of obstructive pulmonary disease in tunnel workers. *Thorax*, **55**, 277–82.

Van Eerdewegh, P., Little, R. D., Dupuis, J. *et al.* (2002). Association of the *ADAM33* gene with asthma and bronchial hyperresponsiveness. *Nature*, **418**, 426–30.

von Mutius, E., Fritzsch, C., Weiland, S. K., Roll, G. and Magnussen, H. (1992). Prevalence of asthma and allergic disorders among children in united Germany: a descriptive comparison. *Bmj*, **305**, 1395–9.

von Mutius, E., Martinez, F. D., Fritzsch, C. *et al.* (1994). Skin test reactivity and number of siblings. *Bmj*, **308**, 692–5.

Walley, A. J., Chavanas, S., Moffatt, M. F. *et al.* (2001). Gene polymorphism in Netherton and common atopic disease. *Nat Genet*, **29**, 175–8.

Wang, X., Wypij, D., Gold, D. R. *et al.* (1994). A longitudinal study of the effects of parental smoking on pulmonary function in children 6–18 years. *Am J Respir Crit Care Med*, **149**, 1420–5.

Wickens, K. L., Crane, J., Kemp, T. J. *et al.* (1999). Family size, infections, and asthma prevalence in New Zealand children. *Epidemiology*, **10**, 699–705.

Wise, R. A. (1997). Changing smoking patterns and mortality from chronic obstructive pulmonary disease. *Prev Med*, **26**, 418–21.

Witt, C., Stuckey, M. S., Woolcock, A. J. and Dawkins, R. L. (1986). Positive allergy prick tests associated with bronchial histamine responsiveness in an unselected population. *J Allergy Clin Immunol*, **77**, 698–702.

Woolcock, A. J. (1996). Asthma – disease of a modern lifestyle. *Med J Aust*, **165**, 358–9.

Woolcock, A. J., Peat, J. K., Salome, C. M. *et al.* (1987). Prevalence of bronchial hyperresponsiveness and asthma in a rural adult population. *Thorax*, **42**, 361–8.

Yamada, N., Yamaya, M., Okinaga, S. *et al.* (2000). Microsatellite polymorphism in the heme oxygenase-1 gene promoter is associated with susceptibility to emphysema. *Am J Hum Genet*, **66**, 187–95.

Yemaneberhan, H., Bekele, Z., Venn, A. *et al.* (1997). Prevalence of wheeze and asthma and relation to atopy in urban and rural Ethiopia. *Lancet*, **350**, 85–90.

Yim, J. J., Park, G. Y., Lee, C. T. *et al.* (2000). Genetic susceptibility to chronic obstructive pulmonary disease in Koreans: combined analysis of polymorphic genotypes for microsomal epoxide hydrolase and glutathione S-transferase M1 and T1. *Thorax*, **55**, 121–5.

Yoshikawa, M., Hiyama, K., Ishioka, S. *et al.* (2000). Microsomal epoxide hydrolase genotypes and chronic obstructive pulmonary disease in Japanese. *Int J Mol Med*, **5**, 49–53.

27

Skeletal disorders

Robert A. Colbert

Introduction

Recent years have witnessed major advances in our understanding of the genetic basis for many skeletal disorders. In several instances, particularly with the less common Mendelian diseases, the responsible gene has been mapped, mutations identified, and their functional significance determined, contributing significantly to our understanding of the molecular basis for pathogenesis. For more common disorders where susceptibility is complex, a complete picture of the responsible genetic polymorphisms and how they influence pathogenesis has been more difficult to achieve. In this chapter we focus on our current understanding of the genetic basis and pathogenic mechanisms for disorders of bone homeostasis. For readers interested in the genetic basis of developmental disorders of the skeleton, we recommend a recent review (Kornak and Mundlos, 2003).

Bone resorption and bone formation are ongoing processes in both the developing and mature skeleton. Even during growth, where the balance favors bone formation, bone resorption is necessary to remove calcified cartilage prior to the formation of mature bone. In the adult skeleton, there continues to be a dynamic balance between these processes that serves both metabolic and mechanical needs of the individual. Although bone mass continues to increase during childhood, it peaks between the ages of 25 and 35 years, and then begins a steady decline that becomes most prominent after age 50 years when bone formation does not fully compensate bone loss. The delicate balance between bone formation and bone resorption is maintained largely by the coordinated actions of two cell types, osteoblasts and osteoclasts. Osteoblasts derive from mesenchymal precursors under the influence of a number of soluble factors including LRP5, Wnt, bone morphogenetic proteins (BMPs), sclerostin, and parathyroid hormone (PTH) (Figure 27.1). Other hormonal factors that play a significant role include vitamin D, estrogens, androgens, and corticosteroids. Osteoblasts synthesize and secrete a collagen-based organic bone matrix, control mineralization, and have an important role in influencing osteoclast function.

Osteoclasts are hematopoietic in origin, arising from the monocyte/macrophage lineage. Osteoclast precursors are stimulated to proliferate by macrophage-colony stimulating factor (M-CSF) secreted locally by osteoblasts. Osteoblasts also express receptor activator of nuclear factor kappa-B (NF-κB) ligand (RANKL) which binds to its receptor, RANK, and stimulates osteoclast differentiation into multi-nucleated cells capable of attaching to, and resorbing, bone (Figure 27.1). Mature osteoclasts express carbonic anhydrase II (CAII), which they utilize to generate intracellular protons (H^+). Protons and chloride (Cl^-) ions are transported into an extracellular space formed between the osteoclast border and the bone surface, thus creating a highly acidic environment. The low pH is critical for bone demineralization,

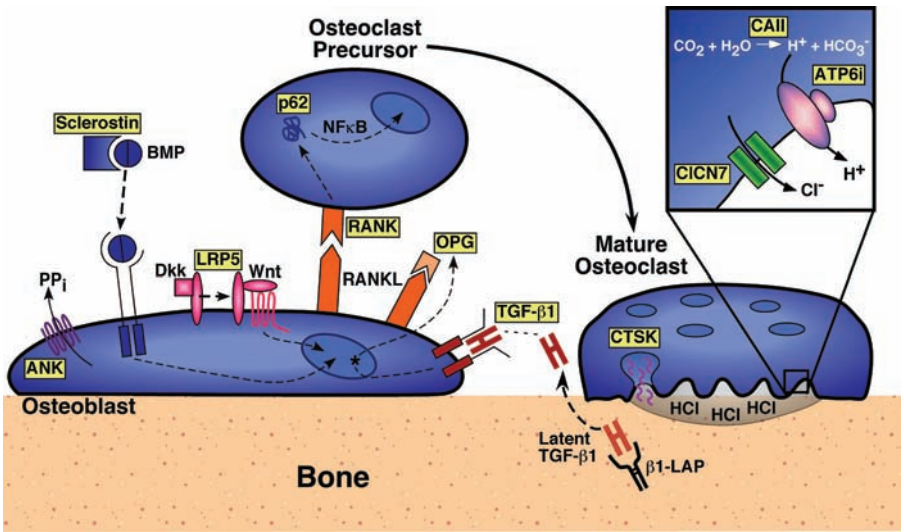

Figure 27.1 Schematic representation of the relationship between osteoblasts and osteoclasts in bone remodeling. Gene products highlighted in yellow boxes are those for which mutations have been shown to have clinically significant effects on bone formation. Heavy dashed arrows signify movement, while lighter dashed arrows imply signaling pathways where details have been omitted for clarity. Mutations in genes encoding the ATP6i subunit of osteoclast H^+-ATPase, the CLC-7 chloride channel (ClCN7), carbonic anhydrase II (CAII), cathepsin K (CTSK), sclerostin, and the pyrophosphate (PPi) transporter ANKH, have all been associated with high bone mass phenotypes. Mutations in genes encoding the osteoclast receptor for RANK ligand (RANK), sequestosome-1 (p62), osteoprotegerin (OPG), and transforming growth factor beta-1 (TGFβ-1) have all been associated with diseases of aberrant bone formation. Mutations in the *LRP5* gene have been associated with both high and low bone mass diseases, depending on the mutation. ANKH mutations have been found in patients with crystal deposition disorders (chondrocalcinosis) as well as aberrant bone formation (craniometaphyseal dysplasia). (Adapted from (Janssens and Van Hul, 2002)).

which is accompanied by enzymatic degradation of the organic matrix by cathepsin K (CTSK) and other proteases. CTSK is an abundant lysosomal enzyme that functions at low pH to degrade several matrix components, but is particularly important for the initial cleavage of type I collagen, rendering it susceptible to other proteases.

Transforming growth factor-β1 (TGF-β1) is also an important regulator of bone homeostasis. Bone degradation releases TGF-β1 which is stored in a latent matrix-embedded form, bound to its latency-associated peptide (β1-LAP). The release of active TGF-β1 from β1-LAP involves the action of proteases including plasmin and thrombospondin. TGF-β1 has a modulating effect on bone resorption by increasing osteoprotegerin (OPG) expression and causing a decrease in RANKL on osteoblasts, which in turn reduces osteoclast activation. In addition to suppressing bone resorption, TGF-β1 also stimulates osteoblast proliferation and differentiation thus promoting bone formation. There are also reports of TGF-β1 stimulating osteoclast formation and bone resorption in in vitro systems. These apparently opposing actions underscore the complexity of regulation, and suggest the process that predominates in vivo may depend on other regulatory factors.

Proper regulation and function of osteoclasts and osteoblasts is critical for homeostatic balance between bone resorption and bone formation. It also allows the individual to temporarily alter the balance to respond to metabolic and/or

mechanical needs. As we will discuss in this chapter, this balance is altered in several genetic disorders, sometimes resulting in low bone mass (osteoporosis), high bone mass (osteopetrosis), or enhanced turnover with dysregulated bone synthesis.

Low bone mass

Osteoporosis

The most common disease of bone, osteoporosis, is characterized by low bone mineral density (BMD), deterioration of bone architecture, and enhanced fragility with increased risk of fracture. Osteoporosis is defined as BMD more than 2.5 standard deviations below the mean established for young adults, while a less severe reduction in BMD (−1 to −2.5 standard deviations) is referred to as osteopenia. Based on this definition, it is estimated that osteoporosis affects 50% of women 80 years of age or older, and nearly 200 million women worldwide. Fractures related to osteoporosis result in significant morbidity and increased mortality. Osteoporosis can be classified as primary or secondary depending on the etiology. Causes of secondary osteoporosis are multiple, and include endocrine dysfunction, medications, immobilization, inflammatory disorders, and congenital forms such as osteogenesis imperfecta. Osteoporosis associated with inflammatory arthritis will be discussed later in this chapter. Readers interested in more detail on the epidemiology and etiologies of secondary osteoporosis should refer to a comprehensive rheumatology or endocrinology textbook.

Genetic differences are estimated to account for over 70% of the variability in BMD (Liu *et al.*, 2003; Ralston, 2003; 2005), and 50–80% of the variability in femoral neck geometry and bone turnover (Arden *et al.*, 1996; Garnero *et al.*, 1996). In contrast, the heritability of fracture itself is relatively low (25–35%), underscoring the importance of fall-related factors such as vision, balance and strength in determining the ''complete'' phenotype of osteoporosis-related fracture (Kannus *et al.*, 1999). Many candidate genes whose products are involved in bone matrix formation and turnover have been investigated for associations with BMD. Table 27.1 provides a summary of genes for which significant associations have been found. One of the first candidate genes studied was the vitamin D receptor (*VDR*) because of its role in calcium metabolism and bone cell function. A meta-analysis incorporating results from several studies revealed only a modest effect on peak bone mass (Cooper and Umbach, 1996). More recently additional *VDR* polymorphisms have been identified, but there are conflicting results as to whether these are associated with osteoporosis. A polymorphism in the first intron of *COLIA1*, which encodes collagen type 1(α), results in increased affinity for the Sp1 transcription factor, and increased expression of the mRNA (Grant *et al.*, 1996; Mann *et al.*, 2001). This results in an altered ratio of α1 to α2 polypeptides in the collagen protein, and a reduction in bone strength. Additional *COLIA1* promoter polymorphisms that are in linkage disequilibrium with the Sp1 binding site polymorphism may also be of functional significance (Garcia-Giralt *et al.*, 2002). The *LRP5* gene is linked to BMD in the lumbar spine, hip, and femoral neck, and single nucleotide polymorphisms (SNPs) have been associated with differences in BMD in several populations encompassing multiple ethnicities, suggesting a robust effect (Ferrari *et al.*, 2004; Koay *et al.*, 2004; Koller *et al.*, 2005; Mizuguchi *et al.*, 2004). *LRP5* encodes the low-density lipoprotein (LDL) receptor-related protein 5 (*LRP5*), which is a co-receptor for Wnt signaling (Figure 27.1) and is important for osteoblast proliferation and development. Interestingly, different mutations in *LRP5* are responsible for an autosomal recessive form of osteoporosis (see below), while others linked to high-bone mass disorders will be discussed in the next section. Other candidate genes listed in Table 27.1 such as those encoding the estrogen receptor, TGF-β1, and IL-6, are included because at least one study has suggested an association with osteoporosis.

Table 27.1. Genes contributing to low bone mass

Disorder	MIM[a] number	Gene/ protein	Chromosomal locus	Phenotype	Inheritance	Pathogenic mechanism
Osteoporosis (primary)	166710				Non-Mendelian	
		VDR/ vitamin D receptor	12q12-q14	Low BMD		
		COLIA1/ type I collagen	17q21-q22	Low BMD, Fx, FNG		Abnormal α1/α2 collagen ratio with weaker bone
		ESRA/ Estrogen receptor-α	6q25.1	Low BMD		
		TGFB1/ TGF-β1	19q13	Low BMD, Fx		
		IL6/IL-6		Low BMD, increased turnover		
		LRP5/LDL receptor-related protein 5	11q13.4	Low BMD		Altered Wnt signaling
Osteoporosis-pseudoglioma syndrome (OPPG)	259770	*LRP5*/LDL receptor-related protein 5	11q13.4	Low BMD; OI with ocular involvement	AR	Reduced Wnt signaling

[a] Online Mendelian Inheritance in Man (OMIM) www.ncbi.nlm.nih.gov/Omim/

TGFB1, Transforming Growth Factor-β1; *IL6*, Interleukin (IL)-6; *LRP5*, Low density lipoprotein receptor-related protein 5; BMD, bone mineral density; FNG, femoral neck geometry; Fx, fracture; OI, osteogenesis imperfecta; AR, autosomal recessive.

However, they are not further discussed either because additional studies have failed to confirm their association, or functional consequences of polymorphisms have yet to be found.

Whole genome screens have been used to locate chromosomal regions linked to BMD. Loci that are linked to traits with a continuous distribution are referred to as quantitative trait loci (QTL). One QTL consistently found for BMD is on chromosome 1p36 (Devoto *et al.*, 2001; Wilson *et al.*, 2003). Polymorphisms in two candidate genes in this region have been associated with low BMD; tumor necrosis factor (TNF) receptor family member 1B (*TNFRSF1B*) and methylene tetrahydrofolate reductase (*MTHFR*) (Albagha *et al.*, 2002; Miyao *et al.*, 2000; Spotila *et al.*, 2000; Tasker *et al.*, 2004). Functional significance for these polymorphisms has not been established. There is a QTL for spinal

BMD on chromosome 1q21-q23 (Koller *et al.*, 2000), and a third region containing a gene encoding the type 1 parathyroid hormone receptor (*PTHr1*) that has been associated (Duncan *et al.*, 1999) and linked to BMD (Wilson *et al.*, 2003). Using less stringent selection criteria for low BMD, two potential QTL (*4q36* and *13q33*) that overlap with genomic regions identified in other linkage studies (Niu *et al.*, 1999) have been identified. Many other QTL found in initial whole genome screens have not been replicated, perhaps because polygenic diseases are not as amenable to linkage analysis (Ralston, 2005).

In summary, despite strong evidence supporting the importance of genetic polymorphisms in a large proportion of the variability in BMD, identification of the genes involved continues to be a challenge. Although environmental factors are likely to play a role, they may be relatively common.

Osteoporosis-pseudoglioma syndrome

Osteoporosis-pseudoglioma syndrome (OPPG) is a rare autosomal recessive condition characterized by juvenile-onset osteoporosis with increased fracture risk and skeletal deformities. Children with OPPG are frequently blind at birth due to vitreous opacities. Inactivating mutations in *LRP5* have recently been shown to cause this disorder in homozygous individuals (Gong *et al.*, 2001), while heterozygous carriers exhibit intermediate reductions in bone mass and have an increased fracture risk (Gong *et al.*, 1996). More recently compound heterozygosity for two novel mutations was found in two family members with OPPG (Cheung *et al.*, 2006). *LRP5* is a coreceptor for Wnt (Figure 27.1) and it plays an essential role in osteoblast proliferation and differentiation (Kato *et al.*, 2002). Inactivation of *LRP5* can severely limit osteoblast production and lead to osteoporotic phenotypes. Activating mutations associated with high bone mass will be further discussed below.

Altered homeostasis of bone resorption and bone formation

Collectively, the disorders associated with enhanced bone turnover are relatively common, primarily due to the inclusion of Paget's disease, and are largely genetically determined (Table 27.2). They are primarily diseases of osteoclast overstimulation with the result being increased formation of abnormal bone that is often deformed and structurally weak.

Paget's disease of bone

Paget's disease of bone (PDB) is one of the most common chronic skeletal diseases affecting up to 3% of many White populations over the age of 60 years. Its ethnic and geographic distribution is variable, with a high prevalence in Whites from the United Kingdom, Australia, North America and Western Europe. PDB is characterized by focal areas of increased bone resorption and formation, leading to deformity and/or enlargement. The axial skeleton (pelvis, lumbar and thoracic spine, and sacrum) is most frequently involved, followed by the femur, skull, and tibia. The newly formed bone in pagetic lesions is disorganized, frequently resulting in bowing and increased fracture. Bony overgrowth in the skull may lead to nerve entrapment, headache, and deafness.

Genetic causes of PDB are heterogeneous. Many families display autosomal dominant transmission, while in other cases only 15–40% of patients have an affected first-degree relative (Morales-Piga *et al.*, 1995; Siris *et al.*, 1991), or it can occur sporadically. Whole genome screens have revealed susceptibility loci on chromosomes 6p21.3 (*PDB1*), 18q21-q22 (*PDB2*), 5q35-qter (*PDB3*), 5q31 (*PDB4*), 2q36 (*PDB5*), 10p13 (*PDB6*), and 18q23 (*PDB7*) (Cody *et al.*, 1997; Haslam *et al.*, 1998; Hocking *et al.*, 2001; Laurin *et al.*, 2001; 2002). Of these regions, the most is known about *PDB2* and *PDB3*. Certain forms of PDB that present under the age of 20 years are due to mutations in RANK, the osteoclast receptor for RANKL (Figure 27.1).

Table 27.2. Genes causing increased bone turnover

Disorder	MIM[a] number	Gene/protein	Chromosomal locus	Phenotype	Inheritance	Pathogenic mechanism
Increased bone resorption and formation						
Paget's disease	602080 (167250)	*TNFRSF11A*/ RANK	18q22.1	High bone turnover, formation of abnormal bone, increased fracture risk	AD	Increased NF-κB activation due to constitutive activation of RANK
		SQSTM1 or *p62*/p62	5q35-qter	Same	AD	Altered NF-κB activation
Juvenile Paget's disease (Idiopathic or hereditary hyperphosphatasia)	239000	*TNFRSF11B*/ OPG	8q24.2	Early onset, more severe form of adult Paget's	AR	Unopposed osteoclast activation due to loss or dysfunction of OPG
Familial expansile osteolysis	174810	*TNFRSF11A*/ RANK	18q22.1		AD	Increased NF-κB activation due to constitutive activation of RANK
Camurati–Engelman disease	131300	*TGFB1*/ TGF-β1	19q13.1		AD	Enhanced TGFβ stimulation of osteoblasts

[a] Online Mendelian Inheritance in Man (OMIM) www.ncbi.nlm.nih.gov/Omim/
Camurati–Engelman disease is also known as progressive diaphyseal dysplasia.
TNFRSF, TNF receptor superfamily; *SQSTM1*, sequestosome 1; RANK, receptor activator of NF-κB; TGFB. Transforming growth factor beta; AD, autosomal dominant; AR, autosomal recessive.

However, RANK mutations are not responsible for most cases of familial or sporadic PDB (Sparks *et al.*, 2001; Wuyts *et al.*, 2001), although they were reported in one Japanese family (Hughes *et al.*, 2000). Interestingly, the gene responsible for familial expansile osteolysis, a rare disorder that overlaps phenotypically with PDB, has also been mapped to *PDB2* (Hughes *et al.*, 1994).

Mutations in the *SQSTM1/p62* gene located in *PDB3* have been found in patients with PDB (Hocking *et al.*, 2002; Laurin *et al.*, 2002). One missense mutation was found in 16% of sporadic and 46% of familial cases, but not in unaffected individuals. The gene encodes a scaffolding protein (sequestosome 1/p62) that plays an important role in several signal transduction pathways. It contains multiple interaction domains that

link receptor interacting protein 1 (RIP) and TNF receptor associated factor 6 (TRAF6) with atypical protein kinase C (aPKC) isoforms that can activate an inhibitor of kappa B kinase (IKK) leading to NF-κB activation (Geetha and Wooten, 2002; Sanz *et al.*, 1999; 2000). Importantly, TRAF6 interacts with RANK expressed on osteoclasts, and is essential for RANK-induced NF-κB activation, which in turn is important in osteoclast formation and differentiation (Figure 27.1) (Darnay *et al.*, 1999). The precise role of p62 in RANK-mediated NF-κB activation, and whether mutations in p62 have functional significance, remains to be determined.

The pathophysiology of PDB has recently been reviewed (Roodman and Windle, 2005). There is evidence suggesting that PDB has a viral etiology (Reddy *et al.*, 2001). Osteoclasts from patients have nuclear and cytoplasmic inclusions resembling paramyxovirus nucleocapsids. Interestingly, paramyxoviral-like nuclear inclusions have been reported in patients with familial expansile osteolysis and occasionally in other bone disorders. Their presence in a very high percentage of PDB osteoclasts has been a consistent finding, and they have even been reported in bone marrow cells. Potential connections between the genetic and viral etiologies have not been established.

Juvenile Paget's disease

Juvenile Paget's disease of bone (JPDB), also called hereditary hyperphosphatasia, is an autosomal recessive disease that begins in infancy or early childhood with pain and bone deformity due to excessive widespread bone remodeling. It is distinct from adult PDB in age of onset and the extensive distribution of lesions. JPDB is much less common with only approximately 50 cases reported worldwide. In JPDB continuous bone turnover impairs growth and remodeling of the entire skeleton, and leads to weak and disorganized bone structure with progressive deformity, osteopenia, and increased fracture risk. Bone histology reveals increases in both osteoclasts and osteoblasts, and typically serum levels of alkaline phosphatase are elevated (hyperphosphatasia). The long bones have widened diaphyses and, along with vertebral and pelvic bones, become progressively deformed. The genetic etiology has been attributed to mutations in the *TNFSFR11B* gene on chromosome 8q24 that encodes osteoprotegerin (OPG) (Figure 27.1). In two families complete OPG deficiency resulted from a 100 kb homozygous deletion at 8q24.2 which included the *TNFRSF11B* gene (Whyte *et al.*, 2002). In a third family, an in-frame deletion resulted in the loss of a conserved amino acid, which rendered it ineffective in suppressing bone resorption *in vitro*, although serum levels of OPG were normal (Cundy *et al.*, 2002). OPG is a decoy receptor for RANKL, and thus can suppress osteoclast activation and bone resorption. Thus, the molecular pathogenesis of JPDB appears to be related to inadequate negative regulation of the RANKL/RANK axis controlling osteoclast activation.

Familial expansile osteolysis

Familial expansile osteolysis (FEO) is characterized by focal regions of increased bone remodeling resembling those seen in adult PDB. There is increased activity of osteoclasts and osteoblasts leading to expansion, deformity and increased risk of fracture. Although PDB and FEO have histologic and phenotypic similarities, they are sufficiently different to be considered separate diseases. FEO begins in the second decade, tends to be more severe, and the bone lesions are predominantly peripheral. Pain is prominent with increased medullary expansion leading to deformity and increased risk of fracture. Many affected individuals have early onset deafness and loss of dentition. FEO has been mapped to the *TNFRSF11A* gene which encodes RANK, where two different tandem duplications in exon 1 encoding the signal peptide have beeen found in several families (Hughes *et al.*, 2000). Signal peptide cleavage is defective, leading to intracellular accumulation of RANK and increased

constitutive NF-κB activation. Expansile skeletal hyperphosphatasia (ESH) is similar to FEO (Whyte *et al.*, 2000), and is associated with a similar tandem duplication in *TNFRSF11A*, suggesting that these diseases are allelic (Whyte and Hughes, 2002). The slight phenotypic differences may be related to other genetic factors.

Camurati–Engelmann disease

Camurati–Engelmann disease (type I) (CED) is an autosomal dominant disorder often presenting with bone pain, particularly in the legs, and non-progressive muscular weakness. It almost always begins before the age of 30 years, and often before age ten years. There is marked thickening of the cortex of bones, both on the periosteal surface and in the medullary canal, usually in a bilaterally symmetric distribution. CED is linked to chromosome 19q13.1-q13.3 (Ghadami *et al.*, 2000) where mutations in *TGFB1*, the gene encoding TGF-β1 have been found (Janssens *et al.*, 2000; Kinoshita *et al.*, 2000). The mutations are predominantly in the latency-associated peptide (β1-LAP) (Janssens *et al.*, 2000; Kinoshita *et al.*, 2000), with one in the leader sequence (Janssens *et al.*, 2003). Several of these mutations appear to facilitate TGF-β1 activation by enhancing its release from β1-LAP (Figure 27.1) (Janssens *et al.*, 2003; McGowan *et al.*, 2003; Saito *et al.*, 2001; Wallace *et al.*, 2004). Although two different mutations reduce the secretion of TGF-β1, the intracellular form appears to be activating, which could account for the similar phenotypes. There is a second form of Camurati–Engelmann (type II) that does not appear to involve TGF-β1 mutations (Nishimura *et al.*, 2002). Its molecular genetic basis remains to be defined.

High bone mass

The diseases with high bone mass, or osteopetrosis, are organized by etiology into decreased resorption and increased formation of bone (Table 27.3). Compared to osteoporosis and Paget's disease, the osteopetroses are far less prevalent, however, deciphering their molecular genetic etiologies has been particularly informative with regard to our understanding of bone homeostasis (Balemans *et al.*, 2005).

Decreased bone resorption

Autosomal recessive osteopetrosis

This form of osteopetrosis, also known as autosomal recessive Albers–Schonberg disease or ''marble'' bone disease, is characterized by macrocephaly, progressive loss of hearing and vision, hepatosplenomegaly, and severe anemia. This form typically presents in early infancy or even during fetal life. Ocular involvement with retinal degeneration occurs in about half by two months of age (Gerritsen *et al.*, 1994). Without bone marrow transplant, the probability of survival to age 6 is only about 30%. The neurologic and hematologic features are thought to result from direct pressure of encroaching bones and bone marrow obliteration, respectively. In a subset of patients loss-of-function mutations in the *TCIRG1/ATP6i* gene, which encodes a regulatory subunit of the H^+-ATPase pump, have been found (Figure 27.1) (Frattini *et al.*, 2000). Loss of ATP6i function results in defective transport of protons by the osteoclast into the extracellular space, and an inability to demineralize bone.

Autosomal dominant osteopetrosis

Autosomal dominant osteopetrosis (also known as autosomal dominant Albers–Schonberg disease) is less severe than the autosomal recessive form. Two phenotypic variants have been described. Type I is characterized by sclerosis of the cranial vault with facial paralysis beginning in the second decade of life. Fractures and osteomyelitis of the mandible are common. In type II, cranial sclerosis is most prominent at the base, and the vertebrae demonstrate endplate thickening resulting in a

Table 27.3. Genes causing diseases of high bone mass

Disorder	MIM[a] number	Gene/protein	Chromosomal locus	Phenotype	Inheritance	Pathogenic mechanism
Decreased bone resorption						
Osteopetrosis, autosomal recessive (Albers–Schonberg disease, AR)	259700	*TCIRG1/ATP6i* ATPase subunit of proton pump	11q12-q13	Increased bone formation and density	AR	Impaired bone demineralization due to impaired acidification
Osteopetrosis, autosomal dominant (Albers–Schonberg disease, AD)	166600	*ClCN7*/Chloride channel 7	16p13.3	Increased bone formation and density	AD	Impaired bone demineralization due to impaired acidification
Osteopetrosis with renal tubular acidosis	259730	*CAII*/Carbonic anhydrase II	8q22	Increased bone formation and density	AR	Impaired bone demineralization due to impaired acidification
Pycnodysostosis	265800	*CTSK*/ Cathepsin K	1q21	Osteosclerosis and bone fragility	AR	Lack of collagen degradation
Increased bone formation						
van Buchem's disease (endosteal hyperostosis)	239100	*SOST*/Sclerostin	17q12-q21	Cortical hyperostosis	AR	Reduced inhibition of BMP signaling
Sclerosteosis	269500	*SOST*/Sclerostin	17q12-q21	Cortical hyperostosis with syndactyly	AR	Reduced inhibition of BMP signaling
High bone mass	601884	*LRP5*/LDL receptor-related protein 5	11q12-q13	High spinal bone mass	AD	Loss of Dkk inhibition
Craniometaphyseal dysplasia	123000	*ANKH*/ANK	5p15.2-p14.1	Sclerosis and hyperostosis of cranial bones	AD; some AR	Overgrowth and sclerosis of craniofacial bones
Chondrocalcinosis	118600	*ANKH*/ANK	5p15	CPPD crystal deposition	AD and sporadic	

[a] Online Mendelian Inheritance in Man (OMIM) http://www.ncbi.nlm.nih.gov/Omim/
CPPD, calcium pyrophosphate dihydrate
AD, autosomal dominant; AR, autosomal recessive.

''sandwich'' appearance. The iliac wings of the pelvis are also sclerotic. Long bones reveal increased cortical thickness and a reduced medullary space without subperiosteal thickening. Many lesions demonstrate a ''bone within a bone'' appearance that is typical of osteopetroses. Fractures are common, and hip osteoarthritis requiring joint replacement may develop. A genome-wide scan in a large family revealed a locus at chromosome 16p13.3 (Benichou *et al.*, 2001), where several missense mutations in the *ClCN7* gene have subsequently been found (Cleiren *et al.*, 2001). *ClCN7* forms a chloride channel located in the ruffled border of the osteoclast (Figure 27.1). This channel transports Cl^- anions into the extracellular space to neutralize the electrical potential that would otherwise be generated by unopposed transport of protons. Complete deletion of mouse *Clcn7* results in a similar osteopetrotic phenotype (Kornak *et al.*, 2001). Mutations in *ClCN7* have also been found in some patients with the autosomal recessive form of osteopetrosis (Cleiren *et al.*, 2001). Different inheritance patterns may occur as a function of the type of mutation and the severity of its effect on the function of the protein. Since the chloride channel functions as a dimer, certain missense mutations that might otherwise have little effect on channel function, may exert a dominant-negative influence.

Osteopetrosis with renal tubular acidosis

Autosomal recessive osteopetrosis accompanied by renal tubular acidosis is also associated with short stature, mental retardation, dental malocclusion and visual impairment. The phenotype usually develops in the first two years of life, and can include cerebral calcifications. Like other forms of osteopetrosis, it is very rare. Early studies on carbonic anhydrase II (CAII) activity in red blood cells from these patients revealed very low levels, with intermediate levels in presumed heterozygotes, leading to the discovery that this enzyme was deficient in many affected patients (Sly *et al.*, 1985). CAII is also expressed in osteoclasts where it catalyzes the formation of protons that are transported via the vacuolar H^+-ATPase into the extracellular space (Figure 27.1). Thus, CAII deficiency also affects osteoclast acidification and bone demineralization. The list of mutations that contribute to CAII dysfunction or deficiency is quite extensive, and includes several missense, splice junction, and nonsense mutations. In general, disease severity correlates positively with mutations resulting in complete loss of the CAII enzyme or its activity.

Pycnodysostosis

Pycnodysostosis is an autosomal recessive sclerosing disorder with a milder phenotype than the autosomal recessive osteopetroses. Patients generally have skull deformities including wide sutures, acroosteolysis of the maxilla and phalanges, osteosclerosis, and bone fragility. Linkage analysis in a kindred with multiple affected individuals led to the identification of a region on chromosome 1q21 harboring the gene (Gelb *et al.*, 1995). Analysis of the positional candidate gene, cathepsin K (*CTSK*), led to the identification of missense and nonsense mutations, as well as an altered stop codon, in several different families (Gelb *et al.*, 1996; Haagerup *et al.*, 2000; Ho *et al.*, 1999). Cathepsin K is a cysteine protease that is expressed in osteoclasts, and is important for proteolytic degradation of organic bone matrix that has been demineralized (Figure 27.1). Although osteoclast numbers are normal in pycnodysostosis, ultrastructural analysis has shown that they have large cytoplasmic vacuoles containing collagen fibrils, consistent with their inability to degrade matrix. In vitro, cathepsin K is able to cleave the intact collagen triple helix, rendering it susceptible to other matrix metalloproteases such as MMP-9. Since acid secretion is unaffected in cathepsin K deficiency, the underlying bone becomes demineralized, yet the organic matrix remains intact.

Increased bone formation

Van Buchem's disease and sclerosteosis

Van Buchem's disease and sclerosteosis have very similar phenotypes with large increases in the amount of bone tissue (Janssens and Van Hul, 2002). Van Buchem's disease usually begins during puberty, with osteosclerosis of the skull, mandible, clavicles, ribs, and long bone diaphyses. In sclerosteosis mainly the skull and mandible are affected, although osteosclerosis and hyperostosis can be more generalized. In both disorders optic atrophy and deafness can occur secondary to bone encroachment on nerves. Facial paralysis is common in sclerosteosis, and may be present at birth or develop shortly thereafter. Sclerosteosis tends to be more severe, and excess height and weight (gigantism) and syndactyly are found in most patients. In a cohort of South Africans followed over a 38-year period, almost half died during the course of the survey (mean age of 33 years), usually from increased intracranial pressure (Hamersma *et al.*, 2003). The gene(s) responsible for van Buchem's disease and sclerosteosis was linked to chromosome 17q12-q21 (Balemans *et al.*, 1999; Van Hul *et al.*, 1998) suggesting that the diseases were allelic. Nonsense and splice variant mutations were found in a novel gene that was named Sclerostin (*SOST*) (Balemans *et al.*, 2001; Brunkow *et al.*, 2001). These mutations were found in several sclerosteosis patients, but not in a family with van Buchem's disease (Balemans *et al.*, 2001), where instead a 52 kb deletion 35 kb downstream of the *SOST* gene has been reported (Balemans *et al.*, 2002; Staehling-Hampton *et al.*, 2002). The deletion was postulated to suppress the expression of sclerostin (Balemans *et al.*, 2002; Staehling-Hampton *et al.*, 2002), although effects on the adjacent downstream gene *MEOX1*, which is also involved in axial skeletal development, have not been ruled out. *SOST* belongs to a family of genes whose products function as decoy receptors that antagonize the actions of BMPs and thus serve as negative regulators of bone formation (Kusu *et al.*, 2003; van Bezooijen *et al.*, 2005; Winkler *et al.*, 2003). Thus, mutations that reduce sclerostin expression or its binding to BMPs, would be expected to increase bone formation and account for the osteosclerotic phenotype. Consistent with this prediction, overexpression of wild type sclerostin in mice results in an opposite phenotype, with decreased osteoblast activity, decreased bone formation, and low bone mass (Winkler *et al.*, 2003).

High bone mass

Investigations of a kindred with high spinal bone mass, but no other syndromic features, revealed a locus on chromosome 11q12-q13 (Johnson *et al.*, 1997) where gain-of-function mutations in *LRP5* have been identified (Little *et al.*, 2002). The same mutation has also been found in a kindred with high bone mass that also exhibited entrapment neuropathies, squaring of the jaw, and hard palate abnormalities (Boyden *et al.*, 2002). Additional *LRP5* mutations (and polymorphisms) have been found in other forms of osteopetrosis including van Buchem's disease (Van Wesenbeeck *et al.*, 2003). The phenotypic differences in disorders seemingly related to *LRP5* raise the possibility that there are other modifying genes and/or environmental factors at work (Little *et al.*, 2002). A candidate for a modifying gene is *TCIRG1/ATP6i*, which is in the same region as *LRP5* and is known to affect bone homeostasis.

LRP5 activity is controlled by an antagonist known as Dickkopf (Dkk) (Figure 27.1). Dkk-family gene products can bind with high affinity to LRPs and prevent Wnt signaling. The *LRP5* mutation associated with high bone mass allows it to escape inhibition by Dkk, and overstimulate osteoblast proliferation and differentiation, thus accounting for the phenotype (Boyden *et al.*, 2002). Osteoporosis pseudoglioma, which has a low bone mass phenotype, is due to inactivating mutations. *LRP5* mutations are of considerable importance not only in terms of understanding the pathophysiology of bone mass, but also since the activating

mutants may prove to be useful as biologic agents for stimulating bone formation (Patel and Karsenty, 2002).

Craniometaphyseal dysplasia

Craniometaphyseal dysplasia (CMD) is characterized by metaphyseal dysplasia with sclerosis and hyperostosis of cranial bones. CMD is usually autosomal dominant, although autosomal recessive inheritance has been reported. The long bones have broadened metaphyses and develop an ''Erlenmeyer flask'' appearance. Hyperostosis of calvarial and facial bones leads to nasal obstruction, compression of the cranial nerves, and hypertelorism. Much of the disability associated with CMD is due to facial nerve palsy and hearing loss secondary to nerve compression. Linkage between autosomal dominant CMD and chromosome 5p15.2-p14.1 has been established in three kindreds (Chandler *et al.*, 2001; Nurnberg *et al.*, 1997). Studies of the positional candidate gene, *ANKH*, have revealed several missense mutations, deletions and insertions (Nurnberg *et al.*, 2001; Reichenberger *et al.*, 2001). *ANKH* encodes a transmembrane protein (ANKH) expressed in osteoblasts that transports inorganic pyrophosphate (PPi) (Figure 27.1). The functional significance of the *ANKH* mutations found in CMD has yet to be established.

Chondrocalcinosis

Chondrocalcinosis is a common cause of intermittent joint pain and arthritis in adults, and is caused by the deposition of calcium pyrophosphate dihydrate (CPPD) crystals in articular cartilage. It also referred to as CPPD deposition disease or familial articular chondrocalcinosis. X-ray examination of joints may reveal a dense narrow band following the epiphyseal contour. Chondrocalcinosis can occur sporadically in association with metabolic disorders (hyperparathyroidism, hemochromatosis and hypothyroidism), or it may be familial, with autosomal dominant inheritance. Although chondrocalcinosis does not involve increased bone formation, it provides an example of an alternate phenotype that can develop from different mutations in a gene that can cause aberrant bone formation. Familial chondrocalcinosis has been linked to chromosome 5p15 (Andrew *et al.*, 1999; Doherty *et al.*, 1991) where mutations in *ANKH* have been found (Pendleton *et al.*, 2002; Williams *et al.*, 2002; 2003). Several mutations appear to result in increased ANKH activity, and thus are predicted to increase synovial fluid PPi levels. In patients with chondrocalcinosis synovial fluid PPi is elevated, which is thought to promote formation of CPPD crystals.

An inactivating mutation in the mouse *Ank* gene causes a phenotype known as murine progressive ankylosis (MPA) (Ho *et al.*, 2000). Loss of *Ank* reduces PPi transport resulting in low extracellular (and high intracellular) PPi concentrations. Extracellular PPi is necessary to prevent another type of crystal (hydroxyapatite) from forming. Thus, the low PPi in the joint space in MPA is thought to promote hydroxyapatite crystal formation leading to loss of motion and eventual ankylosis of joints. Considering the pathophysiology of CPPD crystal formation, this suggests that PPi transport is normally under tight control, and that this may by critical for maintaining crystal-free synovial fluid. Molecular mechanisms underlying *ANKH* mutations that apparently cause increased bone formation in CMD remain to be elucidated.

Altered bone homeostasis in inflammatory arthritis

Rheumatoid arthritis

Rheumatoid arthritis (RA) is an autoimmune disease of unknown etiology that develops in genetically predisposed individuals. It has an estimated prevalence of up to 1% worldwide, although this varies with ethnic and geographic differences. RA is characterized by inflammation

in the synovial lining of the joints that can result in pannus formation and destruction of joint cartilage and bone. Genetic predisposition to autoimmune diseases such as RA is complex. The overall heritability of RA is estimated to be about 60% (MacGregor *et al.*, 2000), indicating that environmental factors are also involved. A significant proportion of the genetic predisposition resides in the major histocompatibility complex (MHC) on chromosome 6p21.3 with alleles of the human leukocyte antigen (HLA) class II *DRB1* gene (Harney *et al.*, 2003). (The role of MHC genes in immune regulation and disease is discussed in detail in Chapters 9, 18) A molecular mechanism explaining the role of HLA class II genes in the pathogenesis of RA has not been established, although numerous ideas have been proposed (Firestein, 2003). Despite their importance, MHC class II genes are only part of the overall genetic susceptibility to RA. Whole genome screens on North American families have revealed evidence for linkage at several loci in addition to the MHC, including *1p13*, *1q43*, *6q21*, *10q21*, *12q12*, *17p13*, and *18q21* (Jawaheer *et al.*, 2003). Importantly, although there is overlap, there are many differences between these results and data from European (Cornelis *et al.*, 1998) and British studies (MacKay *et al.*, 2002). The reason for this is unclear, but may include population differences in susceptibility alleles, and/or the need for larger patient groups and more analytical power. Identification of the responsible genes and polymorphisms, and understanding how they influence a very complex pathogenic process, remains a tremendous challenge.

The effects of RA on bone are local, with erosions and periarticular bone loss, as well as systemic, with generalized osteoporosis (Goldring and Gravallese, 2000; Haugeberg *et al.*, 2003). Osteoporosis is about twice as prevalent in patients with RA as the general population, and is multifactorial, due to inflammatory cytokines, immobilization, and the use of corticosteroids, as well as underlying genetic differences. Osteoclasts play a central role in local and systemic bone loss in RA (Gravallese, 2002; Haugeberg *et al.*, 2003). Many of the cytokines and growth factors produced in synovial tissue promote the differentiation of monocyte/macrophage lineage cells into osteoclasts, including macrophage colony stimulating factor (M-CSF), TNF-α, IL-1, IL-11, IL-17, and parathyroid hormone related peptide (PTHrP). The OPG/RANKL/RANK axis is also critically important for pathologic bone destruction. In addition to its expression on osteoblasts, RANKL is produced by T lymphocytes and synovial fibroblasts (Gravallese *et al.*, 2000; Kong *et al.*, 1999; Takayanagi *et al.*, 2000). Moreover, these cells from RA patients can induce osteoclast formation from peripheral blood mononuclear cells in a RANKL-mediated process (Kotake *et al.*, 2001; Shigeyama *et al.*, 2000; Takayanagi *et al.*, 2000). Observations from animal studies indicate that OPG administration can help to prevent and treat bone erosions in adjuvant induced arthritis (Kong *et al.*, 1999) and in a TNF-α overexpression model (Redlich *et al.*, 2002). Thus, there is strong evidence that the OPG/RANKL/RANK axis is involved in bone resorption in RA pathogenesis, and that targeting this axis may have therapeutic potential in erosive arthritis.

Ankylosing spondylitis

Ankylosing spondylitis (AS) is another inflammatory arthritic disease of unknown etiology that differs in many ways from RA. It belongs to a family of disorders known as spondyloarthritides (SpA) that have a tendency to occur in individuals with the MHC class I allele, *HLA-B27*. From 90 to 95% of individuals with AS carry *HLA-B27* compared to less than 10% of most healthy populations, making this one of the strongest HLA-disease associations. (For interested readers the April, 2004 issue of *Current Molecular Medicine* is devoted to *HLA-B27* and the spondyloarthropathies.) The prevalence of SpA is estimated to be 0.5–0.9%, and thus these disorders represent a significant proportion of inflammatory arthritis. Key clinical features distinguishing these disorders

from RA are involvement of the axial skeleton in addition to peripheral joints, and the association of gastrointestinal inflammation, which is often subclinical. Musculoskeletal manifestations in the SpA are marked by the presence of enthesitis, or inflammation where tendons, ligaments and joint capsules insert on bone (Ball, 1983). In addition to being a primary focus of inflammation, entheses are a site of ossification and eventual ankylosis in these disorders. Relatively little is known about the pathogenesis of new bone formation in the SpA (Zhang *et al.*, 2003). Triggering factors may be biomechanical, microbial, and/or genetic, although the relative importance of each of these has not been established. One proposed mechanism of enthesitis in AS suggests that biomechanical factors induce microtrauma, which then results in bone edema and dysregulated bone formation (McGonagle *et al.*, 2002*b*). Although animal models where *HLA-B27* is over-expressed suggest that this molecule plays a direct role in pathogenesis, and support the involvement of microbes in triggering inflammation, for the most part these rodents do not develop spinal ankylosis (Khare *et al.*, 1995; Rehakova *et al.*, 2000; Taurog *et al.*, 1999). However, recent studies over-expressing additional human β_2-microglobulin (hβ_2m) in rats already transgenic for *HLA-B27/hβ_2m* resulted in greater axial arthritis and ankylosis (Tran *et al.*, 2006), although the mechanism(s) responsible for this effect have not been determined. In humans the spinal predilection for many entheseal lesions has hampered detailed study. Peripheral lesions are clearly inflammatory (McGonagle *et al.*, 2002*a*), but as yet the relative expression of components of the OPG/RANK/RANKL axis or other factors that affect bone formation have not been evaluated.

Genetic predisposition to AS is complex (Sims *et al.*, 2004). The contribution of *HLA-B27* to susceptibility is even greater than *HLA-DR*β*1* alleles in RA. There are families where AS follows *HLA-B27* as an autosomal dominant trait with a penetrance of up to 20%, and the concordance rate in *HLA-B27*-positive monozygotic twins approaches 70%. Nevertheless, most individuals with *HLA-B27* do not develop AS unless other susceptibility alleles are present. It is likely that environmental factors are also important, however, some studies suggest they are ubiquitous (Brown *et al.*, 1997).

Since the discovery over 30 years ago that *HLA-B27* is associated with this disease (Brewerton *et al.*, 1973; Schlosstein *et al.*, 1973), there has been considerable effort aimed at understanding how it might play a role in pathogenesis. Its function as a class I molecule that displays peptides on the cell surface, led to ideas that $CD8^+$ T cell recognition must be involved. It was postulated that *HLA-B27* might have a tendency to select self-peptides that resemble microbial peptides more frequently than other alleles, and thus stimulate the proliferation of autoreactive T cells. Although this idea has not been disproved, the absence of compelling experimental support has led investigators to examine other characteristics of the *HLA-B27* molecule that distinguish it from other MHC class I alleles. One such feature is its tendency to misfold during its assembly with β_2m and peptide in the endoplasmic reticulum (ER). Misfolding involves the formation of *HLA-B27* dimers covalently linked by a disulfide bond and prolonged binding of the chaperone BiP (Dangoria *et al.*, 2002; Mear *et al.*, 1999; Tran *et al.*, 2004). When *HLA-B27* is upregulated, misfolded and BiP-bound complexes accumulate, resulting in ''ER stress'' and activation of the unfolded protein response (UPR) (Turner *et al.*, 2005) (and submitted manuscript). The UPR orchestrates multiple cellular events and gene expression changes that can expand the ER and alleviate ER stress. The impact of the UPR on cells of the immune system has not been fully investigated, but some of the signaling pathways that are affected appear to intersect with other pathways used by the innate immune system and cytokines (Smith *et al.*, submitted). The tendency of some HLA-B27 dimers to be expressed on the cell surface (Allen *et al.*, 1999; Dangoria *et al.*, 2002) has led to suggestions that they may be differentially recognized by $CD4^+$ T cells or other leukocyte receptors, and lead to immune dysregulation (Allen *et al.*, 2001; Boyle *et al.*, 2001; Edwards *et al.*, 2000).

Biochemical evidence for HLA-B27 misfolding in the ER, and the expression of some disulfide-linked complexes on the cell surface is quite strong, but as yet the pathologic significance of misfolding or abnormal cell surface complexes in SpA pathogenesis remains to be established.

The HLA-B27 designation actually represents a family of subtypes that differ by one to several amino acid residues. While most subtypes that are sufficiently prevalent to be assessed are strongly associated with disease, two (B*2706 and B*2709) may be only weakly associated (Sims *et al.*, 2004). Whether this is due to properties of the molecule itself, or genetic differences between the populations where these subtypes are found remains to be established. Expression of these subtypes in rodents may shed light on this issue.

Whole genome screens in AS support the idea that there are several non-MHC loci that contribute to susceptibility (Laval *et al.*, 2001; Zhang *et al.*, 2004). As for many complex genetic diseases, further studies will be required in order to define the genes and ultimately determine polymorphisms and their functional significance. A possible role for the *ANKH* gene in susceptibility to AS has been of particular interest given the progressive ankylosis phenotype associated with *Ank* dysfunction in mice. There has been one report of novel polymorphisms in the 5′-flanking region of the *ANKH* gene that are associated with AS (Tsui *et al.*, 2003), although another study in a British population did not find any relationship (Timms *et al.*, 2003). More recently polymorphisms in *ANKH* were found to be more strongly associated with AS in males (Tsui *et al.*, 2005). Thus, the role of variation in function and/or expression of this interesting candidate gene in AS needs to be further evaluated.

Summary

The skeletal disorders discussed in this chapter cover a large spectrum. Osteoporosis is very common, and the morbidity and mortality associated with this disease has a tremendous impact on health care systems around the world. Diseases like Paget's and the inflammatory arthritides occur in up to a few percent of the population, while the Mendelian disorders tend to be extremely rare, and in some instances have been reported in only a few families. The more common diseases are also more complex genetically, and tend to be influenced to a greater extent by environmental factors. The picture emerging for several of the autoimmune and autoinflammatory diseases is that the environmental contribution, although necessary, is probably ubiquitous and may be unavoidable. Nevertheless, establishing the relative contribution of genetics and environment to disease, and understanding how the various components interact, continues to be one of the foremost challenges of genetic, epidemiology, and biomedical researchers. Despite these challenges, rapid advances in genomics have made it almost routine to hunt down single genes that determine a Mendelian phenotype. Furthermore, understanding the function of the gene product has led to significant insight into the molecular details of both physiologic and pathogenic processes for many skeletal disorders. These insights have underscored the importance of osteoblasts and osteoclasts in bone homoeostasis, and the contribution of the RANKL/OPG/RANK axis in regulating these cells. A significant challenge for the future will be the rational design and delivery of molecules that impact these processes for therapeutic benefit.

Glossary

Hyperostosis — thickening of cortical (compact) bone from deposition of osseous tissue along the subperiosteal and/or endosteal surfaces

Linkage disequilibrium — refers to the statistical correlation between the co-segregation of alleles at more than one locus, usually because the loci are sufficiently close together.

Macrocephaly — enlarged head

Metaphyseal dysplasia — abnormal bone development between the diaphyses (shaft) and epiphyses of long bones
Missense mutation — results in substitution of one amino acid for another.
Nonsense mutation — results in a premature stop codon and truncation of the protein.
Osteosclerosis — increased density of trabecular (spongy) bone
Polydactyly — excess digits
Quantitative trait — a trait having a value with a continuous distribution
Quantitative trait loci (QTL) mapping — a procedure for detecting genes controlling quantitative trait loci and estimating their genetic effects and genome locations.
Syndactyly — fusions of the digits (fingers or toes)

Acknowledgements

The assistance of M. L. DeLay in preparation of Figure 27.1 is greatly appreciated.

REFERENCES

Albagha, O. M., Tasker, P. N., McGuigan, F. E., Reid, D. M. and Ralston, S. H. (2002). Linkage disequilibrium between polymorphisms in the human *TNFRSF1B* gene and their association with bone mass in perimenopausal women. *Hum Mol Genet*, **11**, 2289–95.

Allen, R. L., O'Callaghan, C. A., McMichael, A. J. and Bowness, P. (1999). Cutting edge: HLA-B27 can form a novel beta 2-microglobulin-free heavy chain homodimer structure. *J Immunol*, **162**, 5045–8.

Allen, R. L., Raine, T., Haude, A., Trowsdale, J. and Wilson, M. J. (2001). Leukocyte receptor complex-encoded immunomodulatory receptors show differing specificity for alternative HLA-B27 structures. *J Immunol*, **167**, 5543–7.

Andrew, L. J., Brancolini, V., de la Pena, L. S. *et al.* (1999). Refinement of the chromosome 5p locus for familial calcium pyrophosphate dihydrate deposition disease. *Am J Hum Genet*, **64**, 136–45.

Arden, N. K., Baker, J., Hogg, C., Baan, K. and Spector, T. D. (1996). The heritability of bone mineral density, ultrasound of the calcaneus and hip axis length: a study of postmenopausal twins. *J Bone Miner Res*, **11**, 530–4.

Balemans, W., Ebeling, M., Patel, N. *et al.* (2001). Increased bone density in sclerosteosis is due to the deficiency of a novel secreted protein (SOST). *Hum Mol Genet*, **10**, 537–43.

Balemans, W., Patel, N., Ebeling, M. *et al.* (2002). Identification of a 52 kb deletion downstream of the SOST gene in patients with van Buchem disease. *J Med Genet*, **39**, 91–7.

Balemans, W., Van Den Ende, J., Freire Paes-Alves, A. *et al.* (1999). Localization of the gene for sclerosteosis to the van Buchem disease-gene region on chromosome 17q12-q21. *Am J Hum Genet*, **64**, 1661–9.

Balemans, W., Van Wesenbeeck, L. and Van Hul, W. (2005). A clinical and molecular overview of the human osteopetroses. *Calcif Tissue Int*, **77**, 263–74.

Ball, J. (1983). The enthesopathy of ankylosing spondylitis. *Br J Rheumatol*, **22**, 25–8.

Benichou, O., Cleiren, E., Gram, J. *et al.* (2001). Mapping of autosomal dominant osteopetrosis type II (Albers–Schonberg disease) to chromosome 16p13.3. *Am J Hum Genet*, **69**, 647–54.

Boyden, L. M., Mao, J., Belsky, J. *et al.* (2002). High bone density due to a mutation in LDL-receptor-related protein 5. *N Engl J Med*, **346**, 1513–21.

Boyle, L. H., Goodall, J. C., Opat, S. S. and Gaston, J. S. (2001). The recognition of HLA-B27 by human CD4(+) T lymphocytes. *J Immunol*, **167**, 2619–24.

Brewerton, D. A., Hart, F. D., Nicholls, A. *et al.* (1973). Ankylosing spondylitis and HL-A 27. *Lancet*, **1**, 904–7.

Brown, M. A., Kennedy, L. G., MacGregor, A. J. *et al.* (1997). Susceptibility to ankylosing spondylitis in twins: the role of genes, HLA, and the environment. *Arthritis Rheum*, **40**, 1823–8.

Brunkow, M. E., Gardner, J. C., Van Ness, J. *et al.* (2001). Bone dysplasia sclerosteosis results from loss of the SOST gene product, a novel cystine knot-containing protein. *Am J Hum Genet*, **68**, 577–89.

Chandler, D., Tinschert, S., Lohan, K. *et al.* (2001). Refinement of the chromosome 5p locus for craniometaphyseal dysplasia. *Hum Genet*, **108**, 394–7.

Cheung, W. M., Jin, L. Y., Smith, D. K. *et al.* (2006). A family with osteoporosis pseudoglioma syndrome due to compound heterozygosity of two novel mutations in the LRP5 gene. *Bone*, **39**, 470–6.

Cleiren, E., Benichou, O., Van Hul, E. *et al.* (2001). Albers–Schonberg disease (autosomal dominant osteopetrosis, type II) results from mutations in the ClCN7 chloride channel gene. *Hum Mol Genet*, **10**, 2861–7.

Cody, J. D., Singer, F. R., Roodman, G. D. *et al.* (1997). Genetic linkage of Paget disease of the bone to chromosome 18q. *Am J Hum Genet*, **61**, 1117–22.

Cooper, G. S. and Umbach, D. M. (1996). Are vitamin D receptor polymorphisms associated with bone mineral density? A meta-analysis. *J Bone Miner Res*, **11**, 1841–9.

Cornelis, F., Faure, S., Martinez, M. *et al.* (1998). New susceptibility locus for rheumatoid arthritis suggested by a genome-wide linkage study. *Proc Natl Acad Sci USA*, **95**, 10746–50.

Cundy, T., Hegde, M., Naot, D. *et al.* (2002). A mutation in the gene *TNFRSF11B* encoding osteoprotegerin causes an idiopathic hyperphosphatasia phenotype. *Hum Mol Genet*, **11**, 2119–27.

Dangoria, N. S., DeLay, M. L. *et al.* (2002). HLA-B27 misfolding is associated with aberrant intermolecular disulfide bond formation (dimerization) in the endoplasmic reticulum. *J Biol Chem*, **277**, 23459–68.

Darnay, B. G., Ni, J., Moore, P. A. and Aggarwal, B. B. (1999). Activation of NF-kappaB by RANK requires tumor necrosis factor receptor-associated factor (TRAF) 6 and NF-kappaB-inducing kinase. Identification of a novel TRAF6 interaction motif. *J Biol Chem*, **274**, 7724–31.

Devoto, M., Specchia, C., Li, H. H. *et al.* (2001). Variance component linkage analysis indicates a QTL for femoral neck bone mineral density on chromosome 1p36. *Hum Mol Genet*, **10**, 2447–52.

Doherty, M., Hamilton, E., Henderson, J., Misra, H. and Dixey, J. (1991). Familial chondrocalcinosis due to calcium pyrophosphate dihydrate crystal deposition in English families. *Br J Rheumatol*, **30**, 10–15.

Duncan, E. L., Brown, M. A., Sinsheimer, J. *et al.* (1999). Suggestive linkage of the parathyroid receptor type 1 to osteoporosis. *J Bone Miner Res*, **14**, 1993–9.

Edwards, J. C., Bowness, P. and Archer, J. R. (2000). Jekyll and Hyde: the transformation of HLA–B27. *Immunol Today*, **21**, 256–60.

Ferrari, S. L., Deutsch, S., Choudhury, U. *et al.* (2004). Polymorphisms in the low-density lipoprotein receptor-related protein 5 (*LRP5*) gene are associated with variation in vertebral bone mass, vertebral bone size, and stature in whites. *Am J Hum Genet*, **74**, 866–75.

Firestein, G. S. (2003). Evolving concepts of rheumatoid arthritis. *Nature*, **423**, 356–61.

Frattini, A., Orchard, P. J., Sobacchi, C. *et al.* (2000). Defects in TCIRG1 subunit of the vacuolar proton pump are responsible for a subset of human autosomal recessive osteopetrosis. *Nat Genet*, **25**, 343–6.

Garcia-Giralt, N., Nogues, X., Enjuanes, A. *et al.* (2002). Two new single-nucleotide polymorphisms in the COL1A1 upstream regulatory region and their relationship to bone mineral density. *J Bone Miner Res*, **17**, 384–93.

Garnero, P., Arden, N. K., Griffiths, G., Delmas, P. D. and Spector, T. D. (1996). Genetic influence on bone turnover in postmenopausal twins. *J Clin Endocrinol Metab*, **81**, 140–6.

Geetha, T. and Wooten, M. W. (2002). Structure and functional properties of the ubiquitin binding protein p62. *FEBS Lett*, **512**, 19–24.

Gelb, B. D., Edelson, J. G. and Desnick, R. J. (1995). Linkage of pycnodysostosis to chromosome 1q21 by homozygosity mapping. *Nat Genet*, **10**, 235–7.

Gelb, B. D., Shi, G. P., Chapman, H. A. and Desnick, R. J. (1996). Pycnodysostosis, a lysosomal disease caused by cathepsin K deficiency. *Science*, **273**, 1236–8.

Gerritsen, E. J., Vossen, J. M., van Loo, I. H. *et al.* (1994). Autosomal recessive osteopetrosis: variability of findings at diagnosis and during the natural course. *Pediatrics*, **93**, 247–53.

Ghadami, M., Makita, Y., Yoshida, K. *et al.* (2000). Genetic mapping of the Camurati–Engelmann disease locus to chromosome 19q13.1-q13.3. *Am J Hum Genet*, **66**, 143–7.

Goldring, S. R. and Gravallese, E. M. (2000). Mechanisms of bone loss in inflammatory arthritis: diagnosis and therapeutic implications. *Arthritis Res*, **2**, 33–7.

Gong, Y., Slee, R. B., Fukai, N. *et al.* (2001). LDL receptor-related protein 5 (*LRP5*) affects bone accrual and eye development. *Cell*, **107**, 513–23.

Gong, Y., Vikkula, M., Boon, L. *et al.* (1996). Osteoporosis-pseudoglioma syndrome, a disorder affecting skeletal strength and vision, is assigned to chromosome region 11q12–13. *Am J Hum Genet*, **59**, 146–51.

Grant, S. F., Reid, D. M., Blake, G. *et al.* (1996). Reduced bone density and osteoporosis associated with a polymorphic Sp1 binding site in the collagen type I alpha 1 gene. *Nat Genet*, **14**, 203–5.

Gravallese, E. M. (2002). Bone destruction in arthritis. *Ann Rheum Dis*, **61** Suppl 2, ii84–86.

Gravallese, E. M., Manning, C., Tsay, A. *et al.* (2000). Synovial tissue in rheumatoid arthritis is a source of

osteoclast differentiation factor. *Arthritis Rheum*, **43**, 250–8.

Haagerup, A., Hertz, J. M., Christensen, M. F., Binderup, H. and Kruse, T. A. (2000). Cathepsin K gene mutations and 1q21 haplotypes in AT patients with pycnodysostosis in an outbred population. *Eur J Hum Genet*, **8**, 431–6.

Hamersma, H., Gardner, J. and Beighton, P. (2003). The natural history of sclerosteosis. *Clin Genet*, **63**, 192–7.

Harney, S. M., Newton, J. L. and Wordsworth, B. P. (2003). Molecular genetics of rheumatoid arthritis. *Curr Opin Pharmacol*, **3**, 280–5.

Haslam, S. I., Van Hul, W., Morales-Piga, A. *et al.* (1998). Paget's disease of bone: evidence for a susceptibility locus on chromosome 18q and for genetic heterogeneity. *J Bone Miner Res*, **13**, 911–17.

Haugeberg, G., Orstavik, R. E. and Kvien, T. K. (2003). Effects of rheumatoid arthritis on bone. *Curr Opin Rheumatol*, **15**, 469–75.

Ho, A. M., Johnson, M. D. and Kingsley, D. M. (2000). Role of the mouse ank gene in control of tissue calcification and arthritis. *Science*, **289**, 265–70.

Ho, N., Punturieri, A., Wilkin, D. *et al.* (1999). Mutations of CTSK result in pycnodysostosis via a reduction in cathepsin K protein. *J Bone Miner Res*, **14**, 1649–53.

Hocking, L. J., Herbert, C. A., Nicholls, R. K. *et al.* (2001). Genomewide search in familial Paget disease of bone shows evidence of genetic heterogeneity with candidate loci on chromosomes 2q36, 10p13, and 5q35. *Am J Hum Genet*, **69**, 1055–61.

Hocking, L. J., Lucas, G. J., Daroszewska, A. *et al.* (2002). Domain-specific mutations in sequestosome 1 (*SQSTM1*) cause familial and sporadic Paget's disease. *Hum Mol Genet*, **11**, 2735–9.

Hughes, A. E., Ralston, S. H., Marken, J. *et al.* (2000). Mutations in *TNFRSF11A*, affecting the signal peptide of RANK, cause familial expansile osteolysis. *Nat Genet*, **24**, 45–8.

Hughes, A. E., Shearman, A. M., Weber, J. L. *et al.* (1994). Genetic linkage of familial expansile osteolysis to chromosome 18q. *Hum Mol Genet*, **3**, 359–61.

Janssens, K., Gershoni-Baruch, R., Guanabens, N. *et al.* (2000). Mutations in the gene encoding the latency-associated peptide of TGF-beta 1 cause Camurati–Engelmann disease. *Nat Genet*, **26**, 273–5.

Janssens, K., ten Dijke, P., Ralston, S. H., Bergmann, C. and Van Hul, W. (2003). Transforming growth factor-beta 1 mutations in Camurati–Engelmann disease lead to increased signaling by altering either activation or secretion of the mutant protein. *J Biol Chem*, **278**, 7718–24.

Janssens, K. and Van Hul, W. (2002). Molecular genetics of too much bone. *Hum Mol Genet*, **11**, 2385–93.

Jawaheer, D., Seldin, M. F., Amos, C. I. *et al.* (2003). Screening the genome for rheumatoid arthritis susceptibility genes: a replication study and combined analysis of 512 multicase families. *Arthritis Rheum*, **48**, 906–16.

Johnson, M. L., Gong, G., Kimberling, W. *et al.* (1997). Linkage of a gene causing high bone mass to human chromosome 11 (11q12–13). *Am J Hum Genet*, **60**, 1326–32.

Kannus, P., Palvanen, M., Kaprio, J., Parkkari, J. and Koskenvuo, M. (1999). Genetic factors and osteoporotic fractures in elderly people: prospective 25 year follow up of a nationwide cohort of elderly Finnish twins. *Bmj*, **319**, 1334–7.

Kato, M., Patel, M. S., Levasseur, R. *et al.* (2002). Cbfa1-independent decrease in osteoblast proliferation, osteopenia, and persistent embryonic eye vascularization in mice deficient in Lrp5, a Wnt coreceptor. *J Cell Biol*, **157**, 303–14.

Khare, S. D., Luthra, H. S. and David, C. S. (1995). Spontaneous inflammatory arthritis in *HLA-B27* transgenic mice lacking beta 2-microglobulin: a model of human spondyloarthropathies. *J Exp Med*, **182**, 1153–8.

Kinoshita, A., Saito, T., Tomita, H. *et al.* (2000). Domain-specific mutations in TGFB1 result in Camurati-Engelmann disease. *Nat Genet*, **26**, 19–20.

Koay, M. A., Woon, P. Y., Zhang, Y. *et al.* (2004). Influence of LRP5 polymorphisms on normal variation in BMD. *J Bone Miner Res*, **19**, 1619–27.

Koller, D. L., Econs, M. J., Morin, P. A. *et al.* (2000). Genome screen for QTLs contributing to normal variation in bone mineral density and osteoporosis. *J Clin Endocrinol Metab*, **85**, 3116–20.

Koller, D. L., Ichikawa, S., Johnson, M. L. *et al.* (2005). Contribution of the *LRP5* gene to normal variation in peak BMD in women. *J Bone Miner Res*, **20**, 75–80.

Kong, Y. Y., Feige, U., Sarosi, I. *et al.* (1999). Activated T cells regulate bone loss and joint destruction in adjuvant arthritis through osteoprotegerin ligand. *Nature*, **402**, 304–9.

Kornak, U., Kasper, D., Bosl, M. R. *et al.* (2001). Loss of the ClC-7 chloride channel leads to osteopetrosis in mice and man. *Cell*, **104**, 205–15.

Kornak, U. and Mundlos, S. (2003). Genetic disorders of the skeleton: a developmental approach. *Am J Hum Genet*, **73**, 205–15.

Kotake, S., Udagawa, N., Hakoda, M. *et al.* (2001). Activated human T cells directly induce osteoclastogenesis from human monocytes: possible role of T cells in bone destruction in rheumatoid arthritis patients. *Arthritis Rheum*, **44**, 1003–12.

Kusu, N., Laurikkala, J., Imanishi, M. *et al.* (2003). Sclerostin is a novel secreted osteoclast-derived bone morphogenetic protein antagonist with unique ligand specificity. *J Biol Chem*, **278**, 24113–17.

Laurin, N., Brown, J. P., Lemainque, A. *et al.* (2001). Paget disease of bone: mapping of two loci at *5q35-qter* and *5q31*. *Am J Hum Genet*, **69**, 528–43.

Laurin, N., Brown, J. P., Morissette, J. and Raymond, V. (2002). Recurrent mutation of the gene encoding sequestosome 1 (SQSTM1/p62) in Paget disease of bone. *Am J Hum Genet*, **70**, 1582–8.

Laval, S. H., Timms, A., Edwards, S. *et al.* (2001). Whole-genome screening in ankylosing spondylitis: evidence of non-MHC genetic-susceptibility loci. *Am J Hum Genet*, **68**, 918–26.

Little, R. D., Carulli, J. P., Del Mastro, R. G. *et al.* (2002). A mutation in the LDL receptor-related protein 5 gene results in the autosomal dominant high-bone-mass trait. *Am J Hum Genet*, **70**, 11–19.

Liu, Y. Z., Liu, Y. J., Recker, R. R. and Deng, H. W. (2003). Molecular studies of identification of genes for osteoporosis: the 2002 update. *J Endocrinol*, **177**, 147–96.

MacGregor, A. J., Snieder, H., Rigby, A. S. *et al.* (2000). Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. *Arthritis Rheum*, **43**, 30–7.

MacKay, K., Eyre, S., Myerscough, A. *et al.* (2002). Whole-genome linkage analysis of rheumatoid arthritis susceptibility loci in 252 affected sibling pairs in the United Kingdom. *Arthritis Rheum*, **46**, 632–9.

Mann, V., Hobson, E. E., Li, B. *et al.* (2001). A COL1A1 Sp1 binding site polymorphism predisposes to osteoporotic fracture by affecting bone density and quality. *J Clin Invest*, **107**, 899–907.

McGonagle, D., Marzo-Ortega, H., O'Connor, P. *et al.* (2002*a*). Histological assessment of the early enthesitis lesion in spondyloarthropathy. *Ann Rheum Dis*, **61**, 534–7.

McGonagle, D., Marzo-Ortega, H., O'Connor, P. *et al.* (2002*b*). The role of biomechanical factors and HLA-B27 in magnetic resonance imaging-determined bone changes in plantar fascia enthesopathy. *Arthritis Rheum*, **46**, 489–93.

McGowan, N. W., MacPherson, H., Janssens, K. *et al.* (2003). A mutation affecting the latency-associated peptide of TGFbeta1 in Camurati–Engelmann disease enhances osteoclast formation in vitro. *J Clin Endocrinol Metab*, **88**, 3321–6.

Mear, J. P., Schreiber, K. L., Munz, C. *et al.* (1999). Misfolding of HLA-B27 as a result of its B pocket suggests a novel mechanism for its role in susceptibility to spondyloarthropathies. *J Immunol*, **163**, 6665–70.

Miyao, M., Morita, H., Hosoi, T. *et al.* (2000). Association of methylenetetrahydrofolate reductase (MTHFR) polymorphism with bone mineral density in postmenopausal Japanese women. *Calcif Tissue Int*, **66**, 190–4.

Mizuguchi, T., Furuta, I., Watanabe, Y. *et al.* (2004). LRP5, low-density-lipoprotein-receptor-related protein 5, is a determinant for bone mineral density. *J Hum Genet*, **49**, 80–6.

Morales-Piga, A. A., Rey-Rey, J. S., Corres-Gonzalez, J. *et al.* (1995). Frequency and characteristics of familial aggregation of Paget's disease of bone. *J Bone Miner Res*, **10**, 663–70.

Nishimura, G., Nishimura, H., Tanaka, Y. *et al.* (2002). Camurati-Engelmann disease type II: progressive diaphyseal dysplasia with striations of the bones. *Am J Med Genet*, **107**, 5–11.

Niu, T., Chen, C., Cordell, H. *et al.* (1999). A genome-wide scan for loci linked to forearm bone mineral density. *Hum Genet*, **104**, 226–33.

Nurnberg, P., Thiele, H., Chandler, D. *et al.* (2001). Heterozygous mutations in *ANKH*, the human ortholog of the mouse progressive ankylosis gene, result in craniometaphyseal dysplasia. *Nat Genet*, **28**, 37–41.

Nurnberg, P., Tinschert, S., Mrug, M. *et al.* (1997). The gene for autosomal dominant craniometaphyseal dysplasia maps to chromosome 5p and is distinct from the growth hormone-receptor gene. *Am J Hum Genet*, **61**, 918–23.

Patel, M. S. and Karsenty, G. (2002). Regulation of bone formation and vision by LRP5. *N Engl J Med*, **346**, 1572–4.

Pendleton, A., Johnson, M. D., Hughes, A., Gurley, K. A., Ho, A. M., Doherty, M., Dixey, J., Gillet, P., Loeuille, D., McGrath, R. *et al.* (2002). Mutations in *ANKH* cause chondrocalcinosis. *Am J Hum Genet*, **71**, 933–40.

Ralston, S. H. (2003). Genetic determinants of susceptibility to osteoporosis. *Curr Opin Pharmacol*, **3**, 286–90.

Ralston, S. H. (2005). Genetic determinants of osteoporosis. *Curr Opin Rheumatol*, **17**, 475–9.

Reddy, S. V., Kurihara, N., Menaa, C. and Roodman, G. D. (2001). Paget's disease of bone: a disease of the osteoclast. *Rev Endocr Metab Disord*, **2**, 195–201.

Redlich, K., Hayer, S., Maier, A. *et al.* (2002). Tumor necrosis factor alpha-mediated joint destruction is inhibited by targeting osteoclasts with osteoprotegerin. *Arthritis Rheum*, **46**, 785–92.

Rehakova, Z., Capkova, J., Stepankova, R. *et al.* (2000). Germ-free mice do not develop ankylosing enthesopathy, a spontaneous joint disease. *Hum Immunol*, **61**, 555–8.

Reichenberger, E., Tiziani, V., Watanabe, S. *et al.* (2001). Autosomal dominant craniometaphyseal dysplasia is caused by mutations in the transmembrane protein ANK. *Am J Hum Genet*, **68**, 1321–6.

Roodman, G. D. and Windle, J. J. (2005). Paget disease of bone. *J Clin Invest*, **115**, 200–8.

Saito, T., Kinoshita, A., Yoshiura, K. *et al.* (2001). Domain-specific mutations of a transforming growth factor (TGF)-beta 1 latency-associated peptide cause Camurati–Engelmann disease because of the formation of a constitutively active form of TGF-beta 1. *J Biol Chem*, **276**, 11469–72.

Sanz, L., Diaz-Meco, M. T., Nakano, H. and Moscat, J. (2000). The atypical PKC-interacting protein p62 channels NF-kappaB activation by the IL-1-TRAF6 pathway. *Embo J*, **19**, 1576–86.

Sanz, L., Sanchez, P., Lallena, M. J., Diaz-Meco, M. T. and Moscat, J. (1999). The interaction of p62 with RIP links the atypical PKCs to NF-kappaB activation. *Embo J*, **18**, 3044–53.

Schlosstein, L., Terasaki, P. I., Bluestone, R. and Pearson, C. M. (1973). High association of an HL-A antigen, W27, with ankylosing spondylitis. *N Engl J Med*, **288**, 704–6.

Shigeyama, Y., Pap, T., Kunzler, P. *et al.* (2000). Expression of osteoclast differentiation factor in rheumatoid arthritis. *Arthritis Rheum*, **43**, 2523–30.

Sims, A. M., Wordsworth, B. P. and Brown, M. A. (2004). Genetic susceptibility to ankylosing spondylitis. *Curr Mol Med*, **4**, 13–20.

Siris, E. S., Ottman, R., Flaster, E. and Kelsey, J. L. (1991). Familial aggregation of Paget's disease of bone. *J Bone Miner Res*, **6**, 495–500.

Sly, W. S., Whyte, M. P., Sundaram, V. *et al.* (1985). Carbonic anhydrase II deficiency in 12 families with the autosomal recessive syndrome of osteopetrosis with renal tubular acidosis and cerebral calcification. *N Engl J Med*, **313**, 139–45.

Sparks, A. B., Peterson, S. N., Bell, C. *et al.* (2001). Mutation screening of the *TNFRSF11A* gene encoding receptor activator of NF kappa B (RANK) in familial and sporadic Paget's disease of bone and osteosarcoma. *Calcif Tissue Int*, **68**, 151–5.

Spotila, L. D., Rodriguez, H., Koch, M. *et al.* (2000). Association of a polymorphism in the *TNFR2* gene with low bone mineral density. *J Bone Miner Res*, **15**, 1376–83.

Staehling-Hampton, K., Proll, S., Paeper, B. W. *et al.* (2002). A 52-kb deletion in the SOST-MEOX1 intergenic region on 17q12-q21 is associated with van Buchem disease in the Dutch population. *Am J Med Genet*, **110**, 144–52.

Takayanagi, H., Iizuka, H., Juji, T. *et al.* (2000). Involvement of receptor activator of nuclear factor kappaB ligand/osteoclast differentiation factor in osteoclastogenesis from synoviocytes in rheumatoid arthritis. *Arthritis Rheum*, **43**, 259–69.

Tasker, P. N., Albagha, O. M., Masson, C. B. *et al.* (2004). Association between TNFRSF1B polymorphisms and bone mineral density, bone loss and fracture. *Osteoporos Int*, **15**, 903–8.

Taurog, J. D., Maika, S. D., Satumtira, N. *et al.* (1999). Inflammatory disease in HLA-B27 transgenic rats. *Immunol Rev*, **169**, 209–23.

Timms, A. E., Zhang, Y., Bradbury, L., Wordsworth, B. P. and Brown, M. A. (2003). Investigation of the role of *ANKH* in ankylosing spondylitis. *Arthritis Rheum*, **48**, 2898–902.

Tran, T. M., Dorris, M. L., Satumtira, N. *et al.* (2006). Additional human beta2-microglobulin curbs HLA-B27 misfolding and promotes arthritis and spondylitis without colitis in male HLA-B27-transgenic rats. *Arthritis Rheum*, **54**, 1317–27.

Tran, T. M., Satumtira, N., Dorris, M. L. *et al.* (2004). HLA-B27 in transgenic rats forms disulfide-linked heavy chain oligomers and multimers that bind to the chaperone BiP. *J Immunol*, **172**, 5110–19.

Tsui, F. W., Tsui, H. W., Cheng, E. Y. *et al.* (2003). Novel genetic markers in the 5′-flanking region of ANKH are associated with ankylosing spondylitis. *Arthritis Rheum*, **48**, 791–7.

Tsui, H. W., Inman, R. D., Paterson, A. D., Reveille, J. D. and Tsui, F. W. (2005). *ANKH* variants associated with ankylosing spondylitis: gender differences. *Arthritis Res Ther*, **7**, R513–525.

Turner, M. J., Sowders, D. P., DeLay, M. L. *et al.* (2005). *HLA-B27* misfolding in transgenic rats is associated

with activation of the unfolded protein response. *J Immunol*, **175**, 2438–48.

van Bezooijen, R. L., ten Dijke, P., Papapoulos, S. E. and Lowik, C. W. (2005). SOST/sclerostin, an osteocyte-derived negative regulator of bone formation. *Cytokine Growth Factor Rev*, **16**, 319–27.

Van Hul, W., Balemans, W., Van Hul, E. *et al.* (1998). Van Buchem disease (hyperostosis corticalis generalisata) maps to chromosome 17q12-q21. *Am J Hum Genet*, **62**, 391–9.

Van Wesenbeeck, L., Cleiren, E., Gram, J. *et al.* (2003). Six novel missense mutations in the LDL receptor-related protein 5 (*LRP5*) gene in different conditions with an increased bone density. *Am J Hum Genet*, **72**, 763–71.

Wallace, S. E., Lachman, R. S., Mekikian, P. B. *et al.* (2004). Marked phenotypic variability in progressive diaphyseal dysplasia (Camurati–Engelmann disease): report of a four-generation pedigree, identification of a mutation in TGFB1, and review. *Am J Med Genet A*, **129**, 235–47.

Whyte, M. P. and Hughes, A. E. (2002). Expansile skeletal hyperphosphatasia is caused by a 15-base pair tandem duplication in TNFRSF11A encoding RANK and is allelic to familial expansile osteolysis. *J Bone Miner Res*, **17**, 26–9.

Whyte, M. P., Mills, B. G., Reinus, W. R. *et al.* (2000). Expansile skeletal hyperphosphatasia: a new familial metabolic bone disease. *J Bone Miner Res*, **15**, 2330–44.

Whyte, M. P., Obrecht, S. E., Finnegan, P. M. *et al.* (2002). Osteoprotegerin deficiency and juvenile Paget's disease. *N Engl J Med*, **347**, 175–84.

Williams, C. J., Pendleton, A., Bonavita, G. *et al.* (2003). Mutations in the amino terminus of *ANKH* in two US families with calcium pyrophosphate dihydrate crystal deposition disease. *Arthritis Rheum*, **48**, 2627–31.

Williams, C. J., Zhang, Y., Timms, A. *et al.* (2002). Autosomal dominant familial calcium pyrophosphate dihydrate deposition disease is caused by mutation in the transmembrane protein ANKH. *Am J Hum Genet*, **71**, 985–91.

Wilson, S. G., Reed, P. W., Bansal, A. *et al.* (2003). Comparison of genome screens for two independent cohorts provides replication of suggestive linkage of bone mineral density to 3p21 and 1p36. *Am J Hum Genet*, **72**, 144–55.

Winkler, D. G., Sutherland, M. K., Geoghegan, J. C. *et al.* (2003). Osteocyte control of bone formation via sclerostin, a novel BMP antagonist. *Embo J*, **22**, 6267–76.

Wuyts, W., Van Wesenbeeck, L., Morales-Piga, A. *et al.* (2001). Evaluation of the role of RANK and *OPG* genes in Paget's disease of bone. *Bone*, **28**, 104–7.

Zhang, G., Luo, J., Bruckel, J. *et al.* (2004). Genetic studies in familial ankylosing spondylitis susceptibility. *Arthritis Rheum*, **50**, 2246–54.

Zhang, X., Aubin, J. E. and Inman, R. D. (2003). Molecular and cellular biology of new bone formation: insights into the ankylosis of ankylosing spondylitis. *Curr Opin Rheumatol*, **15**, 387–93.

28

The genetics of common skin diseases

Jonathan Rees

Introduction

Clinical practice suggests that diseases of the skin are common in most world populations. Exactly how common is, for the majority of conditions, impossible to say, as population defined studies for most skin diseases have not been carried out. In the UK, probably typical of many first world countries, about 1:6 consultations in primary care is for a skin disorder, more than for any other disease grouping (McCormick *et al.*, 1996). Of these 1:6 consultations, 25% are for atopic dermatitis, 20% for other forms of dermatitis, 25% for acne, and 10% for psoriasis (McCormick, Fleming and Charlton, 1996). The cumulative incidence of atopic dermatitis and other forms of eczema is perhaps 30% (Williams, 1997). Psoriasis affects 2–3% of the population (Naysmith and Rees, 2003), and skin cancer is the commonest human malignancy in Northern European populations, accounting for half of all cancer cases (Marks, 1995); (Rees, 1998).

Skin disease prevalence data is predictably more limited for third world countries, but point incidence rates of skin disease are in the order of 20–30%, with a greater proportion than in first world countries being due to infectious causes (e.g. fungal disease, infestations, leprosy) (Satimia *et al.*, 1998). Even in the absence of rigorous worldwide studies, it seems reasonable to assume that the frequency of diseases of the skin is higher than that of any other organ.

The inadequacy of the epidemiological data reflects a number of aspects of skin disease that are relevant to genetic studies and the rationale for studying the genetics of skin disease. Skin diseases are rarely fatal and there are a very large number of individual disease entities. Many common disorders are self-limiting or minor, and therapy often surprisingly effective. Care rarely requires admission to hospital (and therefore is not usually recorded in disease registers), is usually managed by primary care physicians, or treatment self-administered by purchase of topical agents from a pharmacist. Diagnosis is usually clinical rather than relying on laboratory investigation. Consequently there is little systematic data on the incidence of most inflammatory skin diseases apart from occasional ad hoc studies.

The situation is somewhat better for skin cancers (Armstrong and Kricker, 2001; Marks, 1995). Melanoma incidence and mortality figures are widely available for many countries, but other forms of skin cancer (non-melanoma skin cancer, NMSC) are frequently not systematically recorded. For example, basal cell carcinoma of the skin, the commonest human malignancy, is not recorded by cancer registries in England. It is important to remember that not all skin tumors (chiefly NMSC) are biopsied, and therefore pathology records under-estimate incidence.

Other general aspects of skin disease merit attention. Because many diseases are so common, and the course and extent so varied, the severity of a condition may be more important clinically than the simple presence or absence of disease. Whether you have atopic dermatitis may be less

important than whether you have bad atopic dermatitis. At best, psoriasis might be a little more troublesome than dandruff, at worst, hospital admission may be required for erythroderma.

There are few functional syndromes of the skin, that is, only rarely is the extent or the nature of the disease such that the physical integrity of the person is endangered (examples of the latter would include widespread blistering, or cardiac failure secondary to erythroderma). The main effect of skin disease is therefore either via symptoms (itch or pain) and impairment of social functioning (i.e. stigma, embarrassment, fear of contagion). These latter characteristics mean that recording of disease incidence can in turn be influenced by a number of social and cultural factors.

Spatiotemporal patterns of common skin diseases

Given the paucity or absence of epidemiological data for most skin diseases, small changes in disease incidence are unlikely to be noticed. Clinical impressions of change in incidence are also confounded by the general increase in provision of healthcare in most developed countries over the last century, leading to a reduced threshold for visiting a physician. In addition, there are organizational differences between countries in what disorders fall within dermatology. For instance, in North America perhaps 50% of dermatology workload falls within what in the UK would be classed as cosmetic surgery or medicine.

Despite the absence of systematic data there are important secular and geographical differences in disease burden that (to this author) seem robust. Atopic dermatitis incidence appears to have increased by a factor of two or more in the last three decades in many developed countries (Williams, 1997). Cancers of the skin in many world populations (chiefly those with pale skin) have also increased two to three-fold over a similar time course (Marks, 1995). By contrast psoriasis incidence seems to have stayed constant. For other common conditions such as acne or skin infections there is no suitable data.

The geographical prevalence of skin disease also varies greatly worldwide. Infection and infestations are especially common in many African countries and other third world nations. By contrast, skin cancer rates are low in those with black skin and highest in those of Northern European ancestry who now live in equatorial regions. Rates of the commonest skin cancers are 15–50 times higher in parts of Australia than in parts of the UK, in populations that largely share the same genetic ancestry (Holme *et al.*, 2000; Rees, 1998; Rees, 2002*b*).

Scope of chapter

Even most full time dermatologists will not see at first hand the majority of skin diseases. Most individual disease entities are rare. Some of these rare disorders, such as epidermolysis bullosa and cancer syndromes such as Gorlin's syndrome or familial melanoma have attracted considerable attention from human geneticists over the last ten years. Many such disorders present with a striking collection of physical signs, rely almost entirely on clinical skills for diagnosis, and are highly penetrant, lending themselves to mapping and subsequent gene identification. Major advances have occurred in our knowledge of the genetics of these rare disorders (van Steensel *et al.*, 1997). For instance, until a decade ago, the biologist of skin was in the ignominious position of not knowing the function of the main class of proteins produced by keratinocytes, namely keratins (van Steensel *et al.*, 1997). A situation akin to a hematologist not knowing the physiological role for hemoglobin in red cells. Subsequently, a large number of rare disorders of the skin have been shown to be due to mutations in keratins, collagens and other structural or signaling proteins in skin. Not withstanding the facility for improved counseling and prenatal diagnosis, the benefit from study of these disorders has been greater in the direction of understanding cell biology rather than in

improving current (and usually inadequate) therapy.

This chapter will focus on disorders in which the inheritance appears genetically complex. Given the nature of clinical advance (Rees, 2002*a*), and the real therapeutic improvements over the last thirty years for many skin diseases, it is not self-evident that studying the genetics of complex skin diseases is the most fruitful research approach to improving dermatological healthcare.

I will take a selective approach dealing with the most common diseases: acne, dermatitis, psoriasis and skin cancer account for over half the dermatology cases seen in primary care and around 80% of the workload of a secondary care dermatologist (Harris *et al.*, 1990). What (little) we know of the genetics of acne, atopic dermatitis and psoriasis will be summarized briefly, and the role of genetic factors in skin cancer dealt with in more detail.

Psoriasis (OMIM 177900)

Psoriasis is a chronic inflammatory dermatosis with a predilection for the extensor surfaces and a widely fluctuating clinical course. It is characterised clinically by raised, scaly red plaques, that vary in size from several millimetres to involving most of the body. In perhaps 15% of cases it is accompanied by a seronegative destructive inflammatory arthropathy (Naysmith and Rees, 2003). Histologically, the lesions are characterised by a dense inflammatory infiltrate comprising polymorphs, lymphocytes and keratinocyte hyperproliferation. The skin does not scar and individual areas of skin return to normality during remission. No intermediate phenotype has been identified and there are no useful blood or systemic markers of disease. A variety of drugs provide clues to etiological mechanisms, with a range of immunosuppressives being effective (cyclosporine, methotrexate, corticosteroids) as well as many other drugs with unknown mechanisms of action (anthralin, ultraviolet radiation, retinoids), and topical vitamin D (Naysmith and Rees, 2003).

A number of surveys of psoriasis have been undertaken. In most populations studied the incidence is around 2–3% with the peak of onset of disease in early adult life and a second peak in old age. The former group are more likely to have a positive family history, joint disease and particular HLA associations. The sex incidence is equal. Incidence rates appear much lower in some Asian communities (Japan, China) and American Indians by perhaps a tenfold factor although no standardized surveys between countries have been carried out (Barker, 2001). There are no known secular trends and there is little convincing evidence for any causal factors in the environment with the exception of a widespread belief that stress exacerbates the disease (but with no robust evidence to substantiate the belief); streptococcal sore throats may precipitate an episode of psoriasis particularly in childhood and early adult life; and a modest association in some studies with alcohol excess, although the direction of causality is unclear, and confounding with other aspects of lifestyle may account for this association (Naldi *et al.*, 2001).

Psoriasis is one of the most heritable chronic diseases of man. Although it does not follow a Mendelian pattern, empirical studies has shown marked familiality. Scandinavian data show the lifetime risk of psoriasis of 0.04, 0.28 and 0.65 if no parent, one parent or both parents, have psoriasis. If there is already one child affected the corresponding risks are 0.24, 0.51, and 0.83 (Swanbeck *et al.*, 1997).

Several population based twins studies have been conducted. Duffy *et al.* in a study of a population based twin register in Australia reported a casewise concordance of 35% in monzygotic twins compared with 12% in dizygotic twins, giving a heritability of 80% (Duffy *et al.*, 1993). In another population study, from Scandinavia, MZ and DZ concordance rates were higher at 0.56 and 0.14, with a heritability of ~90% (Brandrup *et al.*, 1982). In neither of these studies

were all subjects examined directly by a physician, although in the present author's view this is unlikely to have seriously skewed the results. The difference in absolute rates between these studies has been attributed, plausibly, to the greater degree of sunshine in Australia, as ultraviolet radiation improves psoriasis and is used therapeutically. Other studies are available (reviewed by Duffy (Duffy *et al.*, 1993)) but their sampling was based on clinic visits rather than population based ascertainment. Some studies have also suggested that the course of disease is more similar in MZ than in DZ pairs (Brandrup *et al.*, 1982; Farber *et al.*, 1974).

In recent years a number of linkage and association studies have been reported on psoriasis. These originally included studies of a small number of families, and more recently have involved sib pair studies of many hundred families. A number of chromosomal regions have been implicated: 1p35–p34, 1q21, 3q21, 4q13, 4q21, 4q34, 6p21.3, 8q24, 10q22–q23, 11p13, 14q31–q32, 16q12–q13, 17q24–q25, 19p13. A minority of these loci have been confirmed in independent studies (e.g. 6p21.3, 14q31–q32, 17q24–q25 (reviewed by (Bowcock and Barker, 2003)).

It has been known for a long time that psoriasis is associated with particular HLA groups with Cw6 standing out in most Western populations. It is stated that 50% of the heritability of psoriasis is accounted for by this region on chromosome 6p21.3 (Barker, 2001) although current opinion favours the causative gene not being an HLA gene but another, close to the MHC cluster. The nature of the gene is, as yet, unknown. Some of the chromosomal regions overlap with those implicated in atopic dermatitis (e.g. 1q21, 3q21, 17q24–q25) and some with other inflammatory diseases such as rheumatoid arthritis. As yet, the responsible genes at any of the loci implicated in psoriasis have not been identified.

Finally, various association studies for candidate genes (such as TNFα) have been published (Reich *et al.*, 2002). Published risk ratios are modest and confirmatory studies required.

Acne (acne vulgaris, OMIM 604324)

Acne is an inflammatory dermatosis, the genetics of which have not been well studied. Acne is characterized by the development of inflammatory papules, comedones and pustules, leading on to possible disfiguring scarring. At a minimal level of severity, acne is ubiquitous in adolescence, and only a minority of cases will develop scarring or have a clinical course running into the second quarter of life.

Three etiological factors are known: sebum excretion from sebaceous glands, infection, and hair follicle duct obstruction (most sebaceous glands empty into the hair follicle duct). The main determinant of sebum excretion rate is hormonal, principally androgens. Therapies rely on targeting each of the etiological factors: reducing sebum excretion (retinoids, estrogens), infection (antibiotics) and duct obstruction (keratolytics).

No secular trends or geographical differences in incidence for the principal clinical variants of acne are known. Family studies of acne are difficult to interpret because the peak prevalence of the disease is in adolescence, recall of disease imperfect, and the disease course is altered by widely available and effective therapies.

Goulden and colleagues in a UK population provided evidence of clustering of acne within families with a fourfold increase in the chance of a relative of a proband case having acne compared with the relative of a control subject (Goulden *et al.*, 1999). Three twin studies have been reported (one in abstract form only). Bataille in a UK population-based twin survey of middle-aged females (based on recall of earlier disease) reported polychoric correlations of 0.82 and 0.40 for MZ and DZ twins respectively, with a quoted heritability of 81% (Bataille *et al.*, 2002). The result are in keeping with preliminary results (in abstract) from Kirk and colleagues in Australia (Kirk *et al.*, 2004) and the earlier findings of Friedman in the US (Friedman, 1984).

Sebum excretion could be considered an intermediate phenotype and its endocrine physiology is well understood. There is only one study looking at the genetics of sebum excretion. Forty pairs of twins were examined, and differences in sebum excretion rates were greater between DZ than MZ twin pairs, but the statistical analysis was superficial (Walton *et al.*, 1988).

Atopic dermatitis (OMIM 603165)

Atopic dermatitis (AD) is a clinical syndrome characterized by an itchy rash with a variety of morphological cutaneous features that change with age, in association with a positive family history and concomitant presence of other atopic diseases (atopic asthma, hay fever, and occasionally urticaria) (Williams, 1997). The atopic immunological state is characterized by a propensity to develop type 1 IgE mediated responses in response to certain antigens, but the cutaneous immunopathology of atopic dermatitis is characterized by the presence of a T cell and inflammatory cell infiltrate resembling the pattern seen in type IV hypersensitivity reactions (rather than the type 1-like response seen in urticaria). The onset of the rash is typically in early life, peaking at age four years and tending to improve with age, although a large proportion of subjects may develop other forms of eczema later in life (Williams, 1997). Drawing the boundary between mild atopic dermatitis and normality is difficult. Patients may have the symptoms and signs only when provoked by environmental factors, and symptoms may often be mild and patients unaware of the physical signs. Various standardized diagnostic criteria have been proposed but have not been used in most genetic studies (Williams, 1997). Therapy is suppressive with the mainstay being emollients (the skin appears dry) and immunomodulators such as corticosteroids or calcineurin inhibitors and antibacterials for episodes of secondary infection.

The incidence of atopic dermatitis, like other atopic disorders such as asthma and allergy to peanuts appears to have increase threefold over the last 40 years (Williams, 2000). The disease is more common in many developed populations and is affected by sib position: sibs with a higher birth order have a lower rate of disease (Williams, 2000).

That atopic dermatitis runs in families along with the other major atopic diseases was recognized in the original description of atopy (Larsen, 2004). Patients with severe atopic dermatitis have a parent affected in between one-third and two-thirds of cases and in families with one or more first-degree relatives affected, 25–45% of children have atopic dermatitis (Larsen, 2004).

The best population-based twin studies come from Schultz Larsen and colleagues in Denmark; studies which have been repeated over time and provide additional information on secular trends (Larsen, 1993; 2004). They report an increase in the cumulative incidence (up to age 7 years) of atopic dermatitis from 0.03 for a 1960–4 birth cohort to 0.12 for the 1975–9 birth cohort. Pairwise concordances for the most recent cohorts were 0.75 and 0.25 for MZ and DZ twins respectively. They additionally have shown that the risk of AD if the dizygotic partner has the disease is equal to the risk seen in non-twin siblings. In reviewing these results, Larsen comments that the concordance rates are much higher in these studies than in those conducted in Sweden earlier in the century although differences in study design render conclusions based on these differences problematic (Larsen, 1993; 2004). Finally some studies suggest that the risk of developing atopic dermatitis is greater when the mother rather than the father is affected. This could be due to imprinting but other explanations in terms of more shared environment and effects on the fetus are also possible (see Larsen, 2004).

Sib and other family studies have identified a number of loci, some of which show overlap with those implicated in psoriasis. The genes underpinning the linkage are not known. Special mention should be made of the gene underlying Netherton's syndrome (OMIM 256500, *SPINK5*).

Netherton's syndrome is an autosomal recessive disorder characterised by a characteristic rash, hair abnormalities, and the atopic state. It is due to a number of mutations in the *SPINK5* gene that encodes a serine protease inhibitor. Recent studies have shown that a particular variant of the *SPINK5* gene is associated with atopy and atopic dermatitis (Walley *et al.*, 2001) (see Chapter 26).

Studies of serum IgE may be viewed as an intermediate phenotype and twin studies suggest that the heritability of IgE in adults is 0.59 and 0.79 in children (Bazarel *et al.*, 1974).

Skin cancers

There are two main types of skin cancer: melanoma, arising from the neural crest derived melanocytes; and basal cell carcinoma and squamous cell carcinoma derived from keratinocytes (the latter two referred to as non-melanoma skin cancer, NMSC) (Rees, 2002*b*). This broad distinction into two types is useful: NMSC is extremely common and case-fatality extremely low (<0.5%); melanoma by contrast is less common but with a case fatality of ~20% (Rees, 1998).

Melanomas usually present as pigmented lesions with a darkening and increase in size of a pre-existing melanocytic nevus (a benign collection of immature melanocytes) or de novo, as a lesion resembling a mole that is growing and getting darker. Melanoma may metastasize early and fatality is usually due to secondary spread. Because therapy of the tumor once spread has occurred is largely palliative, emphasis has been on primary prevention, early detection and excision of malignant lesions, and identification of high risk groups (Rees, 1998).

Basal cell and squamous cell cancers are keratinocyte-derived tumors that usually present as skin-colored and scaly or crusted lesions on sun-exposed sites. They are rarely fatal. Basal cell carcinomas rarely metastasize, and in 99% of cases can be easily treated using either destructive surgical approaches or more recently pharmacological agents topically applied. Squamous cell carcinomas of the skin can metastasize and are usually treated by excision with or without radiotherapy. Their biological behaviour and clinical impact is far less than for other keratinocyte derived tumors on non sun exposed sites such as the cervix or the oral mucosa.

The main environmental determinant of both melanoma and NMSC is ultraviolet radiation (UVR) exposure (although since the body-site distribution and age incidence of the various tumor types differs, the exact relation between the pattern of UVR exposure and tumor incidence, or other host factors, are also likely to be important). The main genetic determinant of skin cancer is the presence or absence of a susceptible phenotype, the main determinant of a susceptible phenotype being the degree of melanin pigmentation of the skin. The genetics of skin cancer, at least on a global scale, is therefore in large part the genetics of human pigmentation.

Human pigmentation

The color of skin is principally due to the presence of the two main chromophores: melanin and hemoglobin. Differences in skin colour between people (in the resting state) are chiefly due to differences in the amount of melanin produced by melanocytes.

Melanin is a complex mixture of polymers of at least two main classes. Eumelanins are brown or black nitrogenous pigments, insoluble in all solvents, which arise by oxidative polymerization of 5,6 dihydroxyindoles derived from tyrosine via dopaquinone. Pheomelanins are alkali soluble, yellow to red pigments, contain sulphur in addition to nitrogen, and arise by oxidative polymerization of cysteinyldopas via benzothiazine intermediates (Prota *et al.*, 1998). A number of gene products have been identified as being involved in the biosynthesis of melanin including tyrosinase and a family of tyrosinase-related proteins.

Melanin strongly absorbs electromagnetic radiation across both the visible and ultraviolet

spectrum. Maximum absorption is in the shorter wavelengths where DNA damage from UVR is maximal. Melanin is produced within endosome-like structures (melanosomes) within melanocytes and passed to neighbouring keratinocytes. Simple histological staining of keratinocytes show little crescents of melanin on top on the nuclei of keratinocytes, particularly in the basal layer of the epidermis. The importance of melanin in protecting against skin cancer is illustrated by the relative absence of most skin cancer in many African populations, and the dramatically raised rates in those who are unable to synthesis melanin in adequate quantities (albinism (Lookingbill *et al.*, 1995)).

Skin pigmentation due to melanin is a highly heritable trait, and skin color was once used to determine the zygosity of twins. Twin studies of a sample of 134 Australian twins give a heritability of 0.83 for skin color measured at 685nm at the inner forearm (Clark *et al.*, 1981). Values for more UVR-exposed sites such as the forehead are much lower and can be accounted for purely by environmental factors. This initially surprising result reflects large variation in UVR exposure and needs to be taken in context. Pigmentation is of two sorts: constitutive and facultative, the former referring to colour in the absence of UVR and the latter, color in response to UVR. In practice, the upper inner forearm receives UVR (as does even the buttock, a site that is to be preferred in such studies) and it seems likely that the figure quoted of 0.83 is likely to be an underestimate. Similarly, the lower heritability quoted for sun-exposed skin needs to be interpreted in the light of high ambient exposures in Australia, and the limited range of skin colors examined. A study of a more diverse genetic group would by definition decrease the environmental influence.

Study of coat colour mutations in the mouse fancy communities has identified a large number of genes important in coat and skin color (Barsh, 1996; Jackson, 1997). Interference with pigmentation can be due to a number of different processes being affected: failure of successful migration of melanocytes to the skin; inadequate biosynthesis of the melanin polymer; or inadequate packaging and transfer of the mature melanosome into the surrounding keratinocyte. Many of the mutations that have been discovered underpin these varied processes (Barsh, 1996; Jackson, 1997).

Although a large number of loci underlying rare Mendelian disorders in man have been identified following previous work in the mouse, until recently and with one exception, the melanocortin 1 receptor (*MC1R*), the importance of these loci to normal physiological variation in pigmentation in man has been unclear. Studies not yet published (Brilliant, M. personal communication) however suggest that many SNPs in, or close to, loci implicated in coat color mutants in mouse influence human pigmentation. Additionally, following the elucidation of the genetic basis of the golden mutant in zebrafish, the human homolog, SLC24A5, appears to be a key determinant of physiological pigmentary variation in man (Lamason *et al.*, 2005). In the section below I discuss in detail what we have learned of the relation between allelic variation at the *MC1R* and human pigmentary phenotypes. However, it is likely that within several years the majority of normal skin and hair colour variation in man will be accounted for in terms of loci that have already been identified.

The *MC1R* and human pigmentation and skin cancer

In 1993, Roger Cone and co-workers showed that a range of murine coat colour mutations were due to mutations of the *MC1R* gene (Robbins *et al.*, 1993). In the mouse (and subsequently a range of other mammals) loss of function recessive mutations at *mc1r* resulted in a decrease in the ratio of eumelanin to pheomelanin, leading to a yellow (or red) coat. By contrast, dominant gain of function mutations lead to an increase in this ratio, leading to dark brown–black hair (Robbins *et al.*, 1993). Why a change in this ratio leads to red hair and yellow hair in different instances is not clear, except that the ratio of eumelanin to pheomelanin

is obviously an ''incomplete'' measure of hair melanin(s) composition.

In 1995 we showed that sequence variation at the *MC1R* was common in man and that particular variants were associated with red hair and ''pale'' skin (Valverde *et al.*, 1995). Such individuals are characterized clinically by a tendency to burn rather than tan in response to UVR, and pale skin on sites protected from the sun (such as the buttock). This clustering of phenotypic characteristics is known to be associated with an elevated risk of most forms of skin cancer (Rees, 2002*c*). Red hair approximates to an autosomal recessive trait although the penetrance depends on which *MC1R* variants are present (typically penetrance is around 0.85). This bold statement ignores some of the subtleties of phenotype: hair color changes with age from childhood through early and late adulthood, and different body sites frequently have different colored hair.

Given that the *MC1R* is a determinant of pigmentary phenotype and that pigmentary phenotype is a risk factor for both melanoma and NMSC it is not surprising that a number of studies have shown that certain *MC1R* alleles are risk factors for skin cancer (Rees, 2003). To date, over 35 variants of the *MC1R* have been identified (Rees, 2003). Functional studies in heterologous transfection systems, rescue of null mice and human association studies have implicated a large number of them as causal in diminishing activity at the *MC1R* (Rees, 2003). There are a number of issues relevant to human pigmentary variability. Not all the mutations are functionally equivalent. Epidemiological studies suggest that they fall into two groups, differing in the degree to which signalling through the receptor is impaired (Sturm *et al.*, 2003). Functional studies in heterologous cells support this with some, such as the *V60L* allele, showing less inhibition of function than the other alleles, such as the *R151C*, *R160W* and *D294H* (Rees, 2003).

In a UK population, over 50% of the population carry variant alleles (Rees, 2003). Given the different activity of particular variants it is clear that much population variation could be explained by the different combinations of variants. In this respect the trait clearly is not Mendelian. Similar issues surround the relation between *MC1R* and hair color. Most studies have lumped hair colors in a simple categorical classification. However recent work suggests that it may be more meaningful to simply relate *MC1R* genotype to either hair colour using spectrophotometry or the ratio of eumelanin to pheomelanin (Naysmith *et al.*, 2004). From this perspective, both melanin (or color) and signalling activity at the *MC1R* are viewed as continuous traits.

Any relation between *MC1R* and melanoma or non melanoma skin cancer is therefore not surprising but rather to be expected: *MC1R* determines the amount of pigmentation, and pigmentation determines the degree of natural protection to carcinogenic UVR. Odds ratios of 2–5 for skin tumors have been recorded with particular *MC1R* alleles and there is a simple additive effect of two alleles (Sturm, 2002).

A number of important issues remain. Some (Bastiaens *et al.*, 2001) authors have argued that the effects of *MC1R* on cancer risk are not mediated simply by effects on pigmentation but that other pathways are operating. The chief reason for this belief is that after statistically adjusting for skin phenotype, an effect of the *MC1R* is still found in case control studies. If this view is correct, then other pathways are mediating the effects of *MC1R* on cancer risk. αMSH, the ligand for *MC1R*, is known in some situations to be a growth factor for melanocytes, and this would lend some credence to this argument. Against this view, epidemiological studies have found that the association between *MC1R* and cancer applies both to melanocytic lesions and keratinocyte lesions. This author's view is that the fact that *MC1R* status is still a significant factor in the regression analyses after skin color has been taken into account is likely to be due to the inadequacy (and misclassification error) of current measures of phenotype.

MC1R and *p16*

Beyond skin color, the next major genetic risk factor for melanoma is the presence of nevi. The number of nevi is in turn a result of both genetic and environmental (sun exposure) factors (Zhu *et al.*, 1999). Kindreds have been identified with a grossly elevated risk of melanoma due to mutations in the *p16* gene, with many of these subjects showing an increased number of atypical nevi. In an elegant study from Australia, Box and colleagues examined the effects of both *Mc1r* and *p16* in the same families (Box *et al.*, 2001). They showed a clear additive effect between the loci on the time of presentation of melanoma with the penetrance of *p16* increasing from 50–85% dependent on *MC1R* status.

Other genetic factors and skin cancer

An increasing number of studies have recently reported on associations between various candidate genes and skin cancer. Most plausible are reports of associations between genes involved in DNA repair and skin cancer. Patients with xeroderma pigmentosum (OMIM 278730), an autosomal recessive disorder characterized by mutation in one of a number of genes involved in nucleotide excision repair, have a grossly elevated risk of both melanoma and NMSC (often presenting in early childhood with multiple tumors). Various studies have recently reported associations between some of the genes implicated in xeroderma pigmentosum and skin cancer (Dybdahl *et al.*, 1999; Tomescu *et al.*, 2001; Winsey *et al.*, 2000), whilst other studies have reported on associations between polymorphisms in pathways that may play a role in oxidative damage control in skin in response to UVR (such as the glutathione transferases) (Lear *et al.*, 1997; Ramsay *et al.*, 2001) or skin immunity (tumour necrosis factor α) (Hajeer *et al.*, 2000). The reported odds ratios for all these candidate gene studies have tended to be modest (<2) and some results have not been confirmed in other populations, and one suspects are likely to be due to chance.

Conclusions

Skin diseases are common and many of the most common skin disorders appear to be amongst the most highly heritable of the common diseases of man. Studies attempting to identify the genes underpinning atopic dermatitis and psoriasis are ongoing and time will tell whether the biological insight offered will translate into new therapies. Any useful role for prediction of disease state would seem most unlikely.

The studies on skin cancer genetics are however of interest to an audience beyond the dermatologist. We probably understand the biology of skin cancer better than for any other common human tumor. We know the chief cause (UVR), we understand the molecular mechanisms of causation (UVB-induced dimer formation), and we understand some of the natural anti-cancer physiological controls, acting either at the cell level (nuclear excision repair enzymes) and at the organ level (pigmentation). We also have a range of Mendelian disorders which highlight the key elements of these various pathways (Rees, 2002*b*). But pretend for a moment that we didn't possess this level of understanding, and, as for many diseases, attempted to define the heritability and to identify disease genes in skin cancer.

Large and well-conducted twin studies have failed to show a need for heritable factors in explaining NMSC risk (Milan *et al.*, 1998). Similarly, large studies have shown a modest heritable contribution to melanoma risk of 18% (Hemminki *et al.*, 2001). Even when a gene likely to play a role in skin cancer was identified (*MC1R*), the large number of SNPs over just 1 kb meant that functional studies and large association studies were necessary, as the majority of SNPs were not associated with disease, and those that were differed in the degree of their quantitative effect. Without the mouse models and physiological

understanding, one might have tended to imagine that genetic factors were not important in skin cancer and to have given up. At the same time, cancer incidence over the last 30 years has doubled or tripled in many populations, differences that we would almost completely attribute to lifestyle changes. Finally, although treatment for melanoma has changed little, attempts to ensure earlier presentation appear to have had limited success. For NMSC, new treatments abound, although they owe little to genetics or even anything but a superficial knowledge of the biology of skin cancer. These observations suggest some moderation in predicting how understanding of the genetic determinants of common diseases will affect clinical practice.

REFERENCES

Armstrong, B.K. and Kricker, A. (2001). The epidemiology of UV induced skin cancer. *Journal of Photochemistry and Photobiology B: Biology*, **63**, 8–18.

Barker, J.N. (2001). Genetic aspects of psoriasis. *Clin Exp Dermatol*, **26**, 321–5.

Barsh, G.S. (1996). The genetics of pigmentation: from fancy genes to complex traits. *Trends in Genetics*, **12**, 299–305.

Bastiaens, M.T., Huurne, J.A., Kielich, C. *et al.* (2001). Melanocortin-1 receptor gene variants determine the risk of nonmelanoma skin cancer independently of fair skin and red hair. *Am J Hum Genet*, **68**, 884–94.

Bataille, V., Snieder, H., MacGregor, A.J., Sasieni, P. and Spector, T.D. (2002). The influence of genetics and environmental factors in the pathogenesis of acne: a twin study of acne in women. *J Invest Dermatol*, **119**, 1317–22.

Bazarel, M., Orgel, H.A. and Hamburger, R.N. (1974). Genetics of IgE and allergy: serum IgE levels in twins. *J Allergy Clin Immunol*, **54**, 288–304.

Bowcock, A.M. and Barker, J.N. (2003). Genetics of psoriasis: the potential impact on new therapies. *J Am Acad Dermatol*, **49**, no. 2 Suppl, S51–S56.

Box, N.F., Duffy, D.L., Chen, W. *et al.* (2001). *MC1R* Genotype Modifies Risk of Melanoma in Families Segregating *CDKN2A* Mutations. *American Journal of Human Genetics*, **69**, 765–73.

Brandrup, F., Holm, N., Grunnet, N., Henningsen, K. and Hansen, H.E. (1982). Psoriasis in monozygotic twins: variations in expression in individuals with identical genetic constitution. *Acta Dermato-Venereologica*, **62**, 229–36.

Clark, P., Stark, A.E., Walsh, R.J., Jardine, R. and Martin, N.G. (1981). A twin study of skin reflectance. *Annals of Human Biology*, **8**, 529–41.

Duffy, D.L., Spelman, L.S. and Martin, N.G. (1993). Psoriasis in Australian twins. *Journal of the American Academy of Dermatology*, **29**, 428–34.

Dybdahl, M., Vogel, U., Frentz, G., Wallin, H. and Nexo, B.A. (1999). Polymorphisms in the DNA repair gene *XPD*: Correlations with risk and age at onset of basal cell carcinoma. *Cancer Epidemiology Biomarkers & Prevention*, **8**, 77–81.

Farber, E.M., Nall, M.L. and Watson, W. (1974). Natural history of psoriasis in 61 twin pairs. *Archives of Dermatology*, **109**, 207–11.

Friedman, G.D. (1984). Twin studies of disease heritability based on medical records: application to acne vulgaris. *Acta Genet. Med Gemellol. (Roma.)*, **33**, 487–95.

Goulden, V., McGeown, C.H. and Cunliffe, W.J. (1999). The familial risk of adult acne: a comparison between first-degree relatives of affected and unaffected individuals. *British Journal of Dermatology*, **141**, 297–300.

Hajeer, A.H., Lear, J.T., Ollier, W.E. *et al.* (2000). Preliminary evidence of an association of tumour necrosis factor microsatellites with increased risk of multiple basal cell carcinomas In Process Citation. *British Journal of Dermatology*, **142**, 441–5.

Harris, D.W., Benton, E.C. and Hunter, J.A. (1990). The changing face of dermatology out-patient referrals in the south-east of Scotland. *British Journal of Dermatology*, **123**, 745–50.

Hemminki, K., Lonnstedt, I., Vaittinen, P. and Lichtenstein, P. (2001). Estimation of genetic and environmental components in colorectal and lung cancer and melanoma. *Genetic Epidemiology*, **20**, 107–16.

Holme, S.A., Malinovszky, K. and Roberts, D.L. (2000). Changing trends in non-melanoma skin cancer in South Wales, 1988–98. *British Journal of Dermatology*, **143**, 1224–9.

Jackson, I.J. (1997). Homologous pigmentation mutations in human, mouse and other model organisms. *Human Molecular Genetics*, **6**, 1613–24.

Kirk, K., Evans, D.M., Farthing, B. and Martin, N.G. (2004). Genetic and environmental influences on acne in adolescent twins. *Twin Research*, **43**, 190.

Lamason, R.L., Mohideen, M.-A.P.K., Mest, J.R. *et al.* (2005). SLC24A5, a Putative cation exchanger, affects pigmentation in zebrafish and humans. *Science*, **310**, 1782–6.

Larsen, F.S. (1993). The epidemiology of atopic dermatitis. In Burr, M.L. ed., *Epidemioogy of clinical allergy*, Vol **31**, pp. 9–28. Basel: Karger.

Larsen, F.S. (2004). Genetic epidemiology of atopic dermatitis. In Williams, H.C. ed., *Atopic dermatitis*, Cambridge, UK: Cambridge University Press, pp. 113–24.

Lear, J.T., Smith, A.G., Bowers, B. *et al.* (1997). Truncal tumor site is associated with high risk of multiple basal cell carcinoma and is influenced by glutathione S-transferase, GSTT1, and cytochrome P450, *CYP1A1* genotypes, and their interaction. *Journal of Investigative Dermatology*, **108**, 519–22.

Lookingbill, D.P., Lookingbill, G.L. and Leppard, B. (1995). Actinic damage and skin cancer in albinos in northern Tanzania: findings in 164 patients enrolled in an outreach skin care program. *Journal of the American Academy of Dermatology*, **32**, 653–8.

Marks, R. (1995). An overview of skin cancers: incidence and causation. *Cancer*, **75** Suppl, 607–12.

McCormick, A., Fleming, D. and Charlton, J. (1996). *Morbidity statistics from general practice: fourth national survey*, 1991–92. London: OPCS.

Milan, T., Kaprio, J., Verkasalo, P.K. *et al.* (1998). Hereditary factors in basal cell carcinoma of the skin: a population-based cohort study in twins, *British Journal of Cancer*, **78**, 1516–20.

Naldi, L., Peli, L., Parazzini, F. and Carrel, C.F. (2001). Family history of psoriasis, stressful life events, and recent infectious disease are risk factors for a first episode of acute guttate psoriasis: results of a case-control study. *Journal of the American Academy of Dermatology*, **44**, 433–8.

Naysmith, L. and Rees, J.L. (2003). Psoriasis and its management. *Journal of the Royal College of Physicians Edinburgh*, **33**, 104–13.

Naysmith, L., Waterston, K., Ha, T. *et al.* (2004). Quantitative measures of the effect of the melanocortin 1 receptor on human pigmentary status. *Journal of Investigative Dermatology*, **122**, 423–8.

Prota, G., d'Ischia, M. and Napolitano, A. (1998). The chemistry of melanins and related metabolites. In Nordlund, J.J. *et al.*, eds., *The pigmentary system: physiology and pathophysiology*, pp. 307–32. New York: Oxford University Press.

Ramsay, H.M., Harden, P.N., Reece, S. *et al.* (2001). Polymorphisms in glutathione S-transferases are associated with altered risk of non-melanoma skin cancer in renal transplant patients. *Journal of Investigative Dermatology*, **117**, 251–5.

Rees, J.L. (1998). Skin cancer. *Journal of the Royal College of Physicians of London*, **31**, 246–50.

Rees, J. (2002*a*). Complex disease and the new clinical sciences. *Science*, **296**, 698–700.

Rees, J.L. (2002*b*). Skin cancer (Including nevoid basal cell carcinoma syndrome). In Vogelstein, B. and Kinzler, K. eds., *The genetic basis of human cancer*, 2nd edn, pp. 529–48. New York: McGraw-Hill.

Rees, J.L. (2002*c*). Why are redheads so susceptible to melanoma? In Newton-Bishop, J. and Gore, M. eds., *Melanoma: critical debates*, pp. 49–60. Oxford: Blackwell Science.

Rees, J.L. (2003). Genetics of hair and skin colour. *Annual Review of Genetics*, **37**, 67–90.

Reich, K., Mossner, R., Konig, I.R. *et al.* (2002). Promoter polymorphisms of the genes encoding tumor necrosis factor-alpha and interleukin-1beta are associated with different subtypes of psoriasis characterized by early and late disease onset. *Journal of Investigative Dermatology*, **118**, 155–63.

Robbins, L.S., Nadeau, J.H., Johnson, K.R. *et al.* (1993). Pigmentation phenotypes of variant extension locus alleles result from point mutations that alter MSH receptor function. *Cell*, **72**, 827–34.

Satimia, F.T., McBride, S.R. and Leppard, B. (1998). Prevalence of skin disease in rural Tanzania and factors influencing the choice of health care, modern or traditional. *Archives of Dermatology*, **134**, 1363–6.

Sturm, R.A. (2002). Skin colour and skin cancer – *MC1R*, the genetic link. *Melanoma Res*, **12**, 405–16.

Sturm, R.A., Duffy, D.L., Box, N.F. *et al.* (2003). The role of melanocortin-1 receptor polymorphism in skin cancer risk phenotypes. *Pigment Cell Res*, **16**, 266–72.

Swanbeck, G., Inerot, A., Martinsson, T. *et al.* (1997). Genetic counselling in psoriasis: empirical data on psoriasis among first-degree relatives of 3095 psoriatic probands. *British Journal of Dermatology*, **137**, 939–42.

Tomescu, D., Kavanagh, G., Ha, T., Campbell, H. and Melton, D.W. (2001). Nucleotide excision repair gene *XPD* polymorphisms and genetic predisposition to melanoma. *Carcinogenesis*, **22**, 403–8.

Valverde, P., Healy, E., Jackson, I., Rees, J.L. and Thody, A.J. (1995). Variants of the melanocyte-stimulating hormone receptor gene are associated with red hair and fair skin in humans. *Nature Genetics*, **11**, 328–30.

van Steensel, M.A., Steijlen, P.M. and Rees, J.L. (1997). Molecular genetic approaches to skin disease: keratins and keratinisation. *Journal of the Royal College of Physicians of London*, **31**, 379–83.

Walley, A.J., Chavanas, S., Moffatt, M.F. *et al.* (2001). Gene polymorphism in Netherton and common atopic disease. *Nat Genet*, **29**, 175–8.

Walton, S., Wyatt, E.H. and Cunliffe, W.J. (1988). Genetic control of sebum excretion and acne–a twin study. See comments. *British Journal of Dermatology*, **118**, 393–6.

Williams, H.C. (1997). Atopic dermatitis. In Grob, J.J. *et al.*, eds., *Epidemiology, causes and prevention of skin diseases*, pp. 227–59. Oxford: Blackwell Science.

Williams, H.C. (2000). The future research agenda. In Williams, H.C. ed., *Atopic dermatitis*, pp. 247–61. Cambridge, UK: Cambridge University Press.

Winsey, S.L., Haldar, N.A., Marsh, H.P. *et al.* (2000). A variant within the DNA repair gene *XRCC3* is associated with the development of melanoma skin cancer. *Cancer Research*, **60**, 5612–16.

Zhu, G., Duffy, D.L., Eldridge, A. *et al.* (1999). A major quantitative-trait locus for mole density is linked to the familial melanoma gene *CDKN2A*: a maximum-likelihood combined linkage and association analysis in twins and their sibs. *American Journal of Human Genetics*, **65**, 483–92.

29

Molecular genetics of Alzheimer's disease and other adult-onset dementias

P. H. St George-Hyslop

Alzheimer's disease (AD), Lewy body variant of Alzheimer disease (LBV), and the fronto-temporal dementias (FTD) are the three commonest causes of adult-onset dementia. These diseases present in mid to late adult life with progressive defects in memory and higher cognitive functions such as performing complex learned motor tasks (apraxias), reasoning etc. In the fronto-temporal dementias, the clinical syndrome can be overshadowed by behavioral disturbances (disinhibition, aggressivity etc.) and speech disturbances (aphasia), which arise from involvement of the frontal neocortex. The FTD symptom complex frequently also includes additional features such as muscle rigidity, tremor, bradykinesia (Parkinsonism), and motor neuron induced muscle weakness (amyotrophy). In contrast, the clinical features of AD and LBV (recent and immediate memory deficits, deficits in praxis, reasoning and judgement etc.) are those stemming from involvement of the temporal lobe, hippocampus, and the parietal association cortices, with lesser involvement of frontal lobes until late in the disease. LBV overlaps with AD, sharing most of the clinical and neuropathological features of AD, but being differentiated by the presence of prominent visual hallucinations, sensitivity to phenothiazine tranquilizers, and the presence of Lewy bodies (α-synuclein containing intraneuronal inclusions) in neocortical neurons. In all three diseases there is prominent loss of neurons in selected cerebral cortical regions (e.g. hippocampus and temporo-parietal neocortices in AD and LBV; frontal neocortices in FTD). In AD and LBV, a second prominent neuropathological feature is the complex, extracellular, fibrillar deposits in the cortex termed senile or amyloid plaques. These plaques contain a number of proteins including apolipoprotein E and α1-anti-chymotrypsin, but the principal protein component is amyloid-β-peptide (Aβ) derived from a longer precursor (β-amyloid precursor protein – *βAPP*). Finally, AD, LBV and FTD are characterized by the presence of neurofibrillary tangles; intraneuronal inclusions composed of phosphorylated forms of tau, a microtubule-associated protein. These intranuclear phosphorylated tau aggregates have been termed neurofibrillary tangles. The overlap between AD and LBV is considerable, not only in their major clinical and neuropathological features (McKeith *et al.*, 2003), but also in their genetic bases (i.e. both show associations with the apolipoprotein E ε4 variant; see below). As a result, the genetics of AD and LBV will be discussed together. A detailed discussion of the clinical and neuropathologic attributes of these diseases can be found in any standard textbook of neurology and neuropathology.

Genetic basis of Alzheimer's disease and Lewy body variant

In the past several years, a number of genetic epidemiology studies have been undertaken on probands with AD and their families. Cumulatively,

these studies (see review in Katzman and Kawas, 1994) strongly argue that the familial aggregation of AD is not due simply to the high frequency of AD in the general population. These studies reveal that the overall lifetime risk for AD in first-degree relatives of AD probands is about 38% by age 85 years (Lautenschlager *et al.*, 1996). These studies also demonstrate that the majority of cases of familially aggregated AD (FAD) probably reflect a complex mode of transmission such as: (1) one or more common independent, but incompletely penetrant, single autosomal gene defects; (2) a multigenic trait; or (3) a mode of transmission in which genetic and environmental factors interact. Nevertheless, there is a small proportion of AD cases (~10%) which appear to be transmitted as pure autosomal dominant Mendelian traits with age-dependent but high penetrance. Molecular genetic studies on pedigrees with the latter type of FAD with molecular genetic tools has led to the discovery of four different genetic loci associated with inherited susceptibility to AD. Molecular genetic studies have identified four different genes associated with inherited susceptibility to AD. It is suspected that at least two, and possibly several additional AD-susceptibility loci remain to be identified because only about 50% of FAD cases are associated with the four FAD loci known to date.

The amyloid precursor protein

The first gene to be identified in association with inherited susceptibility to AD was the amyloid precursor protein gene (*βAPP*). The *βAPP* gene encodes an alternatively spliced transcript which, in its longest isoform, encodes a single transmembrane spanning polypeptide of 770 amino acids (Kang *et al.*, 1987). The *βAPP* precursor protein undergoes a series of endoproteolytic cleavages (Selkoe, 1994). One of these is mediated by a putative membrane-associated α-secretase, which cleaves βAPP_{695} in the middle of the Aβ peptide domain and liberates the extracellular N-terminus of *βAPP*. The identity of α-secretase is not yet known with certainty, although members of the TACE and ADAM-17 disintegrin metalloprotease family are strong candidates.

The other cleavage pathway involves sequential cleavages by β- and γ-secretases, and generates the 40–42 amino acid Aβ peptide. The first cleavage occurs at the beginning of the Aβ domain, and is mediated by β-secretase (BACE 1), a Type 1 transmembrane glycosylated aspartyl protease, resident in post-Golgi membranes and at the cell surface (Vassar *et al.*, 1999). The second set of cleavages occurs at residues +40, 42 by a putative enzyme activity termed γ-secretase (some recent data suggests that there may be a necessary preceding cleavage at residues +49, +50 and 51; termed ε-cleavage). Both γ-cleavage and ε-cleavage require the presence of the presenilin proteins (see below, and both cleavage events occur mainly in a post-Golgi compartment and/or following re-entry of *βAPP* from the cell surface. The N-terminal product of γ/ε-secretase is Aβ, while the C-terminal product is a labile fragment termed amyloid intracellular domain (AICD), which might act as a signal transduction molecule. The γ/ε-secretase cleavage actually generates a mixture of Aβ peptides, some containing 40 residues, the others containing 42 amino acids. Aβ peptides ending at residue 42 or 43 (long-tailed Aβ) are thought to be more fibrillogenic and more neurotoxic than Aβ ending at residue 40, which is the predominant isoform produced during normal metabolism of *βAPP* (Jarrett and Lansbury, 1993; Lorenzo and Yankner, 1994; Pike *et al.*, 1993; Yankner *et al.*, 1990). Little is currently known about the physiological role (if any) of Aβ. Aβ is removed by several pathways, including degradation by neutral endopeptidase cleavage (e.g. by neprilysin [Yasojima *et al.*, 2001]) or by insulin degrading enzyme (IDE) (Qiu *et al.*, 1998).

The function of *βAPP* is currently unknown. Knockout of the murine *βAPP* gene leads only to minor weight loss, decreased locomotor activity, abnormal forelimb motor activity, and non-specific degrees of reactive gliosis in the cortex (Zheng

et al., 1996). In vitro studies in cultured cells suggest that secreted *βAPP* (βAPPs) can function as an autocrine factor stimulating cell proliferation and cell adhesion. Other studies have implied a role for *βAPP* in: (1) signal transduction by association of *βAPP* with heterotrimeric GTP-binding proteins; (2) a receptor for kinesin-1 during the fast axoplasmic transport of vesicles containing BACE and presenilins, and cleavage of APP (Kamal *et al.*, 2001); and/or (3) a signal transduction molecule in a manner similar to the role of the Notch intra-cellular domain, which mediates Notch signaling following binding of delta to Notch at the cell surface during dorsal axis development in embryogenesis (Cao and Sudhof, 2001; Cupers *et al.*, 2001; Sastre *et al.*, 2001).

Several different missense mutations have been discovered in exons 16 and 17 of the *βAPP* gene in families with early-onset AD (http://www.molgen.ua.ac.be/ADMutations/). All of these mutations either alter APP processing and Aβ production, or alter the propensity of Aβ to aggregate into β-sheet amyloid fibrils. Some of the missense mutations in the *βAPP* gene result in the relative (APP_{717}) or absolute ($APP_{670/671}$) over-production of full length Aβ species ending at residue 42. Other mutations cause the over-production of N-terminally truncated species of Aβ ending at residue 42 (APP_{715}); or the production of Aβ species which have increased propensity to assemble into neurotoxic fibrils (APP_{692}, APP_{693}). Multiple molecular mechanisms have been proposed to explain the neurotoxic effects of Aβ (and especially of small fibrillar aggregates called protofibrils). These include inducing apoptosis by direct effects on cell membranes and by indirect effects, such as potentiating effects of excitatory amino acids, oxidative stress, and increases in intracellular calcium and free radicals (Arispe *et al.*, 1993; Mattson *et al.*, 1992; Mattson and Goodman, 1995). However, Aβ may not be the only cytotoxic product of β- and γ-secretase cleavage because the cytoplasmic C-terminal stub (C31-*βAPP*) is also toxic when over-expressed (Lu *et al.*, 2000).

Apolipoprotein E (APOE)

The association of APOE with inherited susceptibility to AD was uncovered by the concurrence of three lines of investigation. Genetic linkage studies in pedigrees with predominantly late-onset, familially aggregated AD provided suggestive evidence ($z = +2.5$) for an AD susceptibility locus on chromosome 19q12-q13 near the *APOE* gene (Pericak-Vance *et al.*, 1991). Second, analysis of proteins from the CSF which were capable of binding the Aβ peptide revealed that one of the proteins was apolipoprotein E (APOE) (Strittmatter *et al.*, 1993*a*). Finally, APOE is a component of the senile plaque of AD.

The *APOE* gene in humans contains three common polymorphisms: ε2 (cysteines at codon 112 and codon 158); ε3 (cysteines at codon 112); and ε3 (arginine at codon 112). Analysis of these polymorphisms in normal control populations and in patients with AD has shown that there is an increase in the frequency of the *ε4* allele in patients with AD (allele frequency in AD is approximately 40%, compared to 15% in normals) (Saunders *et al.*, 1993), and that there is a smaller reduction in the frequency of the *ε2* allele (from 10% to about 2% in AD) (Corder *et al.*, 1994). More significantly, there is a dose-dependent relationship between the number of copies of *ε4*, and the age-of-onset of AD such that ε4/ε4 subjects have an earlier onset than do heterozygous ε4 subjects (Corder *et al.*, 1993). Subjects with an *ε2* allele, on the other hand, have a later onset (Corder *et al.*, 1994). The association between *ε4* and AD has been robustly confirmed in numerous studies and in several different ethnic groups. The association is weaker with advanced age of onset, and the putative protective role of the *ε2* allele is less clear at younger ages of onset (where it may even be associated with a more aggressive course) (Rebeck *et al.*, 1994; van Duijn *et al.*, 1994). This association of APOE *ε4* with AD has been replicated in numerous studies and in numerous ethnic groups with the possible exception of black Americans and American Hispanics, which have

generated conflicting results (Hendrie *et al.*, 1995; Maestre *et al.*, 1995).

Although the association between APOE *ε4* and AD is robust, it is not entirely specific. Observations in patients with head injury (Mayeux *et al.*, 1995; Roses and Saunders, 1995); spontaneous intracerebral hemorrhage (Alberts *et al.*, 1995), and in patients undergoing elective cardiac bypass surgery (Newman *et al.*, 1995), all suggest a poorer outcome for patients with the *ε4* allele. There is a confirmed association between the *ε4* allele and the Lewy body variant of AD (see above) (Olichney *et al.*, 1996).

The mechanism by which the *ε4* allele is associated with an earlier onset of AD, and by which the *ε2* allele is associated with a later onset is unclear. The most obvious hypothesis is that APOE might influence the production, distribution, or clearance of the Aβ peptide. This hypothesis is supported by observations that the genotype at APOE modulates age-of-onset in subjects carrying the *βAPP Val717Ile* mutation (but not the APP_{692} mutation), suggesting a direct biochemical interaction between APOE and *βAPP* (or its metabolic products) (St George-Hyslop *et al.*, 1994). Second, subjects with one or more APOE *ε4* alleles have a higher Aβ peptide plaque burden than do subjects with no *ε4* alleles (Schmechel *et al.*, 1993). In vitro studies suggest that delipidated APOE *ε4* binds Aβ more avidly than APOE *ε3* (Strittmatter *et al.*, 1993*a*; 1993*b*). There is also evidence that both APOE and Aβ may be cleared through the lipoprotein-related (LRP) receptor and that APOE *ε4* and the Aβ peptide may compete for clearance through the LRP receptor (Kounnas *et al.*, 1995). Finally, transgenic mice expressing the βAPP_{V717F} mutation (PDAPP mice) develop profound cerebral Aβ deposition when bred on an $APOE^{+/+}$ background, but have very little Aβ deposition on an $APOE^{-/-}$ background (Bales *et al.*, 1997).

An alternate hypothesis concerns changes in cholesterol metabolism. Both epidemiological and direct experimental evidence in cell culture models suggests that cholesterol metabolism and APP metabolism are functionally intertwined. Specifically, reduction in cellular cholesterol availability results in significant changes in APP trafficking and processing, with the resultant reduction in Aβ formation (Runz *et al.*, 2002). In addition, patients who have taken statins for hypercholesterolemia appear to have a reduced incidence of Alzheimer's disease (Fassbender *et al.*, 2001; Jick *et al.*, 2000; Wolozin *et al.*, 2000).

Finally, there is a good correlation between the degree of clinical dementia and the decrease in synaptic density in AD (Terry *et al.*, 1994), and it has been suggested that APOE may be involved in synaptic plasticity during regeneration and repair, and that the *ε4* allele is less efficient in this role. This is in accord with clinical epidemiological data suggesting that the presence of APOE ε4 is associated with a poorer outcome after a variety of unrelated central nervous system injuries including head injury, stroke, and coronary artery bypass grafting. It has therefore been suggested that the association between APOE *ε4* and AD may not determine whether AD occurs, but rather, the clinico-pathologic response to other causative factors by modulating a variety of effects including Aβ processing and regeneration-repair etc. Indeed, these putative effects of APOE on several different mechanisms need not be mutually exclusive.

Presenilin 1

Genetic mapping studies located a third Alzheimer susceptibility locus (*AD3*) to a region of approximately 10 centiMorgans on the long arm of chromosome 14. The actual disease gene (presenilin 1) was isolated using a positional cloning strategy (Sherrington *et al.*, 1995).

The presenilin 1 (*PS1*) gene, is highly conserved in evolution, being present in *C. elegans* (Levitan and Greenwald, 1995) and *D. melanogaster* (Boulianne *et al.*, 1997). PS1 encodes a polytopic integral membrane protein. Currently favoured models have 8 transmembrane domains, with a large, hydrophilic, acidically charged loop domain between the putative sixth and seventh

transmembrane domains. The PS1 protein is approximately 50 kDa in size and is predominantly located within intracellular membranes in the endoplasmic reticulum, the perinuclear envelope, the Golgi apparatus and at the cell surface as well as in some as yet uncharacterized intracytoplasmic vesicles (De Strooper *et al.*, 1997; Walter *et al.*, 1996). Only very small amounts of the PS1 holoprotein exist within the cell at any given time (Podlisny *et al.*, 1997; Thinakaran *et al.*, 1996). Instead, the holoprotein undergoes endoproteolytic cleavage near residue 290 within the TM6-TM7 loop domain. The biologically active form of the presenilins are the N- and C-terminal fragments (NTF and CTF), which are tightly associated with each other in a high molecular weight, intracellular protein complex (>450 kDa). This protein complex co-elutes with γ-secretase activity in cellular fractionation studies. Several other components of the presenilin complexes have been identified. These include nicastrin (a~110 kDa Type 1 transmembrane glycoprotein (Yu *et al.*, 2000)); APH-1 [a polytopic transmembrane protein that may act as the initial assembly molecule for the complex (Gu *et al.*, 2003)]; and PEN2 (a short hydrophobic protein with two transmembrane domains, but with unknown function).

The presenilins play either a direct catalytic or indirect (facilitatory) role in the proteolytic processing (''regulated intramembranous proteolysis'') of several Type 1 transmembrane proteins including APP, p75, LRP, ErbB4, and notch. Null mutations in the presenilins result in severe developmental defects, which directly arise from defective proteolysis of notch with the subsequent failure to generate the Notch intracellular domain (NICD) (De Strooper *et al.*, 1999; Levitan and Greenwald, 1995; Wong, 1996; Ye *et al.*, 1999). Loss of presenilin function, or loss of nicastrin, APH-1, or PEN-2 have similar effects on notch processing and on the processing of APP. The latter defect causes a profound reduction in the secretion of Aβ and the accumulation of the substrate for γ/ε-cleavage (i.e. C83-APP (α-stubs) and C99 (β-stubs), which are the derivatives of α- and β-secretase cleavage respectively). It is currently unclear whether the presenilins have a direct enzymatic activity (Wolfe *et al.*, 1999), or whether the presenilins are indirectly involved perhaps through modulating trafficking and/or activation of various substrates or other components of the γ-secretase complex. Further details on this controversy are in (Checler, 2001; Wolfe, 2001). To date, more than 128 different mutations have been discovered in the *PS1* gene (http://www.molgen.ua.ac.be/ADMutations/). The majority of these mutations are missense mutations giving rise to the substitution of a single amino acid. A few in-frame splicing, deletion or insertion defects have also been identified (Kwok *et al.*, 1997; Perez-Tur *et al.*, 1995; Sato *et al.*, 1998). However, nonsense mutations resulting in truncated proteins which would cause loss-of-function mutations have yet to be found in AD-affected subjects.

All *PS1* mutations associated with AD increase γ-secretase cleavage of βAPP and preferentially increase the production of toxic long-tailed Aβ peptides ending at residue 42 (Citron *et al.*, 1997; Duff *et al.*, 1996; Martins *et al.*, 1995; Scheuner *et al.*, 1996). However, some investigators believe that like the βAPP mutations, PS1 and PS2 mutations may also cause neurodegeneration by modulating cellular sensitivity to apoptosis induced by a variety of factors, including staurosporine, Aβ peptide, serum withdrawal and so on.

Presenilin 2

During the cloning of the presenilin 1 gene on chromosome 14, an homologous sequence (Presenilin 2) was identified on chromosome 1 (Rogaev *et al.*, 1995). *PS2* encodes a polypeptide whose open reading frame contains 448 amino acids, with substantial sequence similarity to *PS1* (overall identity approximately 60%), and a very similar structural organization. Despite this similarity, *PS1* and *PS2* are likely to have distinct but overlapping functions. For instance, *PS2* does not functionally replace either the APP or Notch processing defects in $PS1^{-/-}$ animals (Herreman

et al., 1999), yet PS2 mutations, like PS1 mutations, increase the secretion of long-tailed Aβ peptides (Citron *et al.*, 1997; Scheuner *et al.*, 1996).

Mutational analyses have uncovered a small number of missense mutations (~9) in the presenilin 2 gene in families segregating early-onset forms of AD (www.molgen.uia.ac.be/ADMutations). The phenotype associated with PS2 mutations is much more variable (Bird *et al.*, 1996; Sherrington *et al.*, 1996). Thus, the vast majority of heterozygous carriers of missense mutations in the *βAPP* and *PS1* genes develop the illness between the ages of 35 and 55 years for PS1 mutations, and between 45 and 65 for βAPP mutations. In contrast, the range of age-of-onset in heterozygous carriers of PS2 mutations is between 40 and 85 years, and there is at least one instance of apparent non-penetrance in an asymptomatic octogenarian transmitting the disease to affected offspring (Bird *et al.*, 1988; 1989; Sherrington *et al.*, 1996). Furthermore, in contrast to APP mutations, the effect of APOE *ε4* on the age-at-onset in PS2 mutations is either absent or less profound. Modifier loci other than APOE probably account for much of this variation.

Other genes for Alzheimer's disease

Several large surveys of patients with familial AD have indicated that the *APP*, *APOE*, *PS1* and *PS2* genes account for only about half of the genetic risk factors for AD. It is therefore likely that there are several additional AD-susceptibility genes. Some of these loci will be associated with additional rare, but highly penetrant defects similar to those seen with mutations in PS1 and APP. Other genes may result in incompletely penetrant autosomal-dominant traits like those associated with PS2. However, it is likely that a significant proportion of the remaining genes will be genes with low-effect sizes similar to APOE, and in which the ultimate phenotype is likely to be influenced by the presence or absence of other genetic and environmental risk factors.

There have been multiple strategies deployed to try to map these additional AD-susceptibility genes. Genetic linkage studies and family-based association analyses have been employed on datasets with pedigrees multiply affected with AD, and have led to the suggestion that there may be additional susceptibility loci in: (1) the pericentromeric region of chromosome 12 (Alzheimer Type 5) (Pericak-Vance *et al.*, 1997; Rogaeva *et al.*, 1998); and (2) the long arm and pericentromeric region of chromosome 10q (Alzheimer Type 6) (Bertram *et al.*, 2000; Ertekin-Taner *et al.*, 2000; Myers *et al.*, 2000). However, to date, the exact genes in these regions that cause susceptibility to AD have not been identified. Weaker evidence, in single studies have also implicated chromosome 20p (Olson *et al.*, 2002); 15q22 (Scott *et al.*, 2003); and 9p (Pericak-Vance *et al.*, 2000). The glutathione S-transferase omega-1 (GSTO1) gene on distal chromosome 10q has also been implicated as a gene modulating age-of-onset in both AD and Parkinson's disease (Li *et al.*, 2003), but this has not yet been replicated in other datasets.

A second strategy to identify AD genes has been to use cohorts of sporadic AD cases and age/sex matched controls in a case:control association design. This has led to a long list of potential candidate genes, most of which have not been robustly replicated. A partial list of candidate genes provisionally identified as putative AD susceptibility loci includes homozygosity for the *AA* allele of an intronic polymorphism in α1-chymotrypsin, *A5* repeat allele of an intronic insertion-deletion polymorphism in the very low density lipoprotein receptor, neutral coding sequence and intronic polymorphisms in low density lipoprotein receptor related protein, homozygosity for intronic polymorphisms in presenilin 1 or presenilin 2, K-variant of butyrylcholinesterase, homozygosity for the Val/Val variant of the Val443Ile polymorphism in bleomycin hydroxylase; IL4 etc. However, most of these candidate genes have not received the same widespread confirmation as did APOE *ε4* when tested in independent but comparable datasets. Interestingly, a number of obvious

candidate genes involved in APP processing, including *BACE* (β-secretase involved in Aβ generation) and neprilysin and insulin degrading enzyme (involved in Aβ degradation) have been screened, but as of this writing have not been widely found to have genetic variants associated with increased risk for AD.

Familial non-specific dementia

A small number of pedigrees have been described which segregate an autosomal dominant adult-onset disease characterized by early-onset dementia with non-specific neuropathological changes of neuronal loss, minor gliosis and some spongiform change (Gydesen *et al.*, 1987). Completely absent are the typical findings of Alzheimer's disease (AD) or Pick's disease. Genetic linkage studies have mapped one such disease locus in one large Danish pedigree to a 12 cM region of chromosome 3 spanning the centromere between the markers D3S1284 and D3S1603 (Brown *et al.*, 1995).

Familial British dementia (FBD) and familial encephalopathy with neuronal dementia with neuroserpin deposits

Recently, two very rare forms of inherited dementia (Familial British dementia and familial encephalopathy with neuroserpin inclusion bodies (FEN1B)) have been described which support the emerging concept that many of the inherited dementias are disorders of protein processing in which there is either intracellular or extracellular toxic accumulation of misfolded/misprocessed proteins, a theme common to all of the diseases discussed here.

Familial British dementia is characterized by spasticity, ataxia, and later progressive dementia accompanied by widespread demyelination, with distinctive perivascular fibrous deposits that are clearly different from the plaques of AD. A T−>A transversion mutation in the stop codon of the *BRI* gene on chr 13 causes the addition of several amino acids at the C-terminus of the BRI protein (Vidal *et al.*, 1999). While the normal function of the BRI protein is unknown, the presence of the extra C-terminal amino acids causes the protein to be misprocessed by a Furin-mediated cleavage. This causes the accumulation of a 34 amino acid C-terminal derivative termed the ABRI peptide that assembles into toxic amyloid deposits through mechanisms which remain to be elucidated (Kim *et al.*, 1999).

Missense mutations in neuroserpin, a neuron-specific serine protease inhibitor (serpin), have been described in two pedigrees with a familial dementia (Davis *et al.*, 1999). In these families with FEN1B the mutant neuroserpin forms typical serpin loop-sheet polymers that assemble into fibrous aggregates which appear as 5−50 μm PAS-positive eosinophilic inclusions called ''Collins bodies'' in the deep layers of the cerebral cortex, subcortical nuclei and substantia nigra.

Fronto-temporal lobe dementia (FTD)

Fronto-temporal dementia is a pleomorphic neurodegenerative illness which typically begins before the age of 65 years. In a minority of cases, the disease is inherited as an autosomal-dominant trait (Lynch *et al.*, 1994; Wszolek *et al.*, 1992). The disease often begins with personality and behavioral changes including disinhibition manifest by alcoholism, hyper-religiosity, hypersexuality, hyperphagia (elements of the Kluver-Bucy syndrome) and stealing. As the disease progressed in these families, further abnormalities in judgment, language and praxia developed. In addition to these cognitive changes, some patients also developed Parkinsonism and amyotrophy. However, presentations with primary progressive aphasia, Parkinsonism, dystonia and/or oculomotor disturbances are not infrequent. Neuropathologically, the illness is typified by fronto-temporal atrophy with severe neuronal loss, spongiform change in the superficial layers

of the neocortex of the temporal lobe and frontal cortex, neuronal loss and gliosis in the substantia nigra and amygdala, and anterior horn cell loss in the spinal cord (Sima *et al.*, 1996). The unifying theme amongst these pedigrees is the predilection for involvement of the frontal and temporal lobes with prominent neuronal loss and gliosis, some spongiform change, and the absence of amyloid or Lewy bodies. In some cases, notably those arising from mutations in the tau gene (see below) the other prominent neuropathological feature is the presence of tau fibrils in neurons and/or glia.

Genetic linkage studies in the subset of FTD cases showing autosomal dominant inheritance revealed that between 10 and 40% of these familial FTD cases showed genetic linkage between a series of DNA markers on chromosome 17 near the *tau* gene (see below). It was also apparent that a significant proportion of autosomal dominant fronto-temporal dementia pedigrees did not show linkage to this region. Genetic linkage studies in these cases subsequently led to the discovery of an additional genetic locus associated with autosomal dominant fronto-temporal dementia with amyotrophic lateral sclerosis on chromosome 9q.

The clinical phenotype for the FTDP-17 cases is, as described above, quite pleomorphic. However, one unifying observation has been that there is pathological deposition of hyper-phosphorylated tau either in neurons or in glia. In contrast, in the chromosome 9 variant, the clinical phenotype tends to have more prominent motor signs, and in those individuals who have a fronto-temporal dementia phenotype. In addition, the neuropathological picture is one of fronto-temporal atrophy with gliosis vacuolar changes, but with a relative paucity of amyloid plaques and neurofibrillary tangles.

Tau mutations in FTDP-17

Sequencing of the tau gene led to the discovery of several different mutations in some families with FTDP-17 (Clark *et al.*, 1998; Hutton *et al.*, 1998; Poorkaj *et al.*, 1998; Spillantini *et al.*, 1997).

Tau is an abundant phosphoprotein that is predominantly expressed in both central and peripheral nervous systems. In brain it is found predominantly in neurons, with lower levels in some glial cells. Within nerve cells, it is found mainly in axons. The major physiological function of Tau is to bind to and stabilize microtubules, and it may therefore play a role in microtubule-dependent axoplasmic transport. Microtubule binding is mediated by three or four microtubule-binding domain repeats (depending on alternative splicing of the tau in RNA).

Tau is a phosphoprotein whose phosphorylation is developmentally regulated. Phosphorylation is thought to be regulated by one or more of several kinases including mitogen-activated protein kinase, glycogen synthase kinase-3, neuronal cdc2-like kinase and stress-activated protein kinases. The phosphorylation state of tau is balanced by the dephosphorylating activity of protein phosphatase 2A (PP2A). Hyperphosphorylation is an invariant feature of the filamentous Tau deposits that characterize a number of neurodegenerative diseases.

Tau mutations causing FTDP-17 have been either: (1) missense and deletion mutations in the coding region; or (2) intronic mutations located close to the 5′-splice site of the intron following exon 10. The tau missense mutations have generally been located in the microtubule-binding repeat region or close to it. Most missense mutations reduce the ability of tau to interact with microtubules, as reflected by a marked reduction in the ability of mutated tau proteins to promote microtubule assembly. Intronic mutations are located at positions +3, +13, +14 and +16, relative to the first nucleotide of the splice-donor site of the intron following exon 10. Secondary structure predictions show that these intronic mutations disrupt an RNA stem-loop structure at the exon 10–5′-intron boundary and lead to increased splicing of exon 10 and increased transcription encoding exon 10. This is reflected by a

net overproduction of four-repeat isoforms. Additional pathogenic mutations may exist in other introns of the *Tau* gene. Thus, a G to A transition at position +33 of the intron following exon 9 has been described in a patient with familial frontotemporal dementia. It disrupts one of several (A/T)GGG repeats that may play a role in the regulation of the alternative splicing of exon 10.

Validation of the hypothesis that over-expression of wild-type *4R-Tau* or over-expression of mutant *3R/4R-Tau* can cause neurodegeneration and intracellular accumulation of fibrillar deposits of tau has been obtained from the analysis of transgenic mice. Over-expression of either wild-type *4R-Tau* or of *P301L* mutant *Tau* transgenes results in neurodegeneration with NFT and spheroids particularly in amygdala, brainstem and cerebellar nuclei, and spinal cord. This was associated with hind limb weakness, amyotrophy, and dystonic posturing (Lewis *et al.*, 2000; Spittaels *et al.*, 1999). Similar features have been observed in a mouse expressing the R406W mutation (M. Ikeda, Gunma University, in preparation).

The mechanism by which tau mutations or over-expression of specific isoforms of tau cause neurodegeneration is unclear. There is some evidence that they alter the binding of tau to microtubules, causing increased amounts of free tau, which then inhibits kinesin-dependent intracellular transport of organelles, neurofilaments and cargo vesicles, including those carrying amyloid precursor protein (APP) (Mandelkow and Mandelkow, 2002). This latter observation, and the prior observation that misprocessing of tau causes neurodegeneration by itself has also led to experiments in which transgenic mice over-expressing mutant human tau transgenes were either crossbred with other transgenic mice expressing mutant APP transgenes (Lewis *et al.*, 2001), or received intracerebral injections of Aβ42 (Gotz *et al.*, 2001). These experiments show a modest exacerbation of the tau pathology by pre-existing Aβ pathology or by exposure to Aβ. This data has been interpreted to support the notion that defects in APP processing in AD may initiate a cascade of other events which themselves add to the neurotoxic cascade (i.e. tau misprocessing) and lead to neurodegeneration. If correct, this brings together the final common pathways for several neurodegenerative diseases.

Practical implications

Predictive genetic testing

The discovery that mutations in specific genes are associated with inheritable susceptibility to late onset dementias, together with the increasing public awareness both of genetics as a cause of these illnesses, and of familial aggregation of AD in particular, lead to the frequent need for physicians to consider the merits of genetic counseling and genetic testing. At the present time, in the absence of clearly effective preventative or curative treatments without significant side effects, the main reason for genetic counseling and testing is to provide information only. However, while such information can be empowering, it obviously also has the potential to be misused to the patient's disadvantage.

There is currently relatively little practical experience with genetic counseling of members of families multiply affected with AD or the other dementias. Consequently, most of the paradigms used for genetic counseling of members of families with dementia are based upon similar paradigms used in the counseling of subjects with Huntington's disease (Copley *et al.*, 1995). The Huntington's disease model is actually quite useful for counseling of members of families with early-onset familial Alzheimer's disease (FAD) associated with mutations in *PS1*, *PS2* or *βAPP* and of FTD associated with mutations in Tau because the age-of-onset is often similar (30–65 years of age) and has a similar pattern of transmission (highly penetrant, age-dependent penetrance, autosomal-dominant segregation). Thus, in members of families with mutations in the *βAPP*, *PS1* and *Tau* genes, it is possible to screen at-risk

family members for the presence of mutations detected in affected individuals, and to counsel these family members based upon the concept that *βAPP*, *PS1* and *Tau* mutations are highly penetrant (approximately 95%) with typical age-of-symptom-onset between 35 and 65 years. PS2 mutations on the other hand have a lower penetrance and a more variable age-of-onset (45–85 years). Using the Huntington's disease paradigm, we and several other groups have had the opportunity to counsel a small number of members of families with *PS1* and *βAPP* mutations without significant problems. We have also shown that screening for PS1 mutations is cost-effective when done on symptomatic cases with a positive family history of AD and onset before the age of 60 years (Rogaeva *et al.*, 2001).

Although use of the Huntington's disease paradigm will probably work well for *PS1*, *tau* and *βAPP* mutations, mutations in these genes are a comparatively rare cause of familial dementias, cumulatively accounting for about 50% of early-onset FAD (which itself accounts for ~5% of all AD) and 10–40% of familial frontotemporal dementia. A much more common clinical experience is the presence of two or three affected family members with late-onset dementia in a small nuclear pedigree. Frequently, in these ''multiplex'' late-onset dementia pedigrees, the disease does not inherit as a classic autosomal-dominant trait. As a result, in any given family, it is frequently unclear whether the multiplex pedigree structure reflects an incompletely penetrant autosomal-dominant trait, or a more complex mode of transmission involving either the additive effects of several genes or the interaction of genes and environment. Empirical counseling of pedigrees of this type is difficult and, in the case of AD where the data are the more robust, must largely be based upon recent epidemiological studies such as those of Lautenschlager *et al.* (1996). While the use of molecular genetic studies would clearly facilitate counseling in such pedigrees, the only locus which has shown robust association with late-onset AD is the apolipoprotein E (*APOE*) gene. Retrospective studies, in autopsy and clinical series, have suggested that the cumulative lifetime risk for AD in subjects homozygous for *APOE ε4* may be as high as 90% by age 90 years. However, even from these retrospective studies, it is apparent that there is a huge variation in the age-of-onset of AD even in subjects who are homozygous for *APOE ε4* (50–90 years). To confound matters further, a small number of limited prospective studies suggest that the *APOE* genotype is a relatively poor predictor of the onset of AD even in high risk groups such as those with age-associated memory loss (Tierney *et al.*, 1996). Consequently, a number of research groups have recommended that the *APOE* genotype not be used for presymptomatic testing (Relkin *et al.*, 1996). The question of whether the *APOE* genotype may be of assistance in establishing the diagnosis of AD in demented patients undergoing diagnostic work-up is currently a matter of some discussion. Most experts agree that while the *APOE* genotype studies might form a part of the diagnostic armamentarium, they are unlikely by themselves to be the only test which should be done even in individuals with classical clinical features of AD (Mayeux *et al.*, 1998). Thus Mayeux and colleagues have shown that a clinical diagnosis of AD has a sensitivity of ~93% and specificity of 55%, whereas the the *APOE ε4* allele confers sensitivity and specificity of 65% and 68%, respectively. The addition of information about the *APOE* genotype increased the overall specificity to 84% in patients who met the clinical criteria for AD, but the sensitivity decreased 61%. However, it is important to note that the patients studied in the report by Mayeux *et al.* were referred to tertiary centres, and it remains to be determined whether genotype would provide similar levels of sensitivity and specificity in more typical clinical settings (Liddell *et al.*, 2001; Mayeux *et al.*, 1998).

Pharmacogenomics

Another role of genetic testing is in the design of a therapeutic program for subjects affected with AD

and other dementias. Given the obvious genetic heterogeneity of AD (and the other dementias), it would not be unreasonable to suspect that this etiologic heterogeneity would be reflected in different biochemical pathogenic pathways and therefore differences in response to specific treatments. It is conceivable, therefore, that some subsets of AD (or the other dementias) might respond better to specific therapeutic agents than other subtypes. For instance, there was provisional evidence that the APOE genotype may be associated with differences in response to the cholinesterase inhibitor Tacrine (Farlow, 1997). However, this doesn't seem to have been born out as a general phenomenon because as additional AChEase inhibitors have been developed, there has been no predictive value for APOE (Aerssens *et al.*, 2001).

It is apparent, from the discussion on the molecular genetics of AD and related disorders, that a significant proportion of familially aggregated dementia cannot be related to polymorphisms or mutations in any of the known genes (*PS1*, *PS2*, *βAPP*, *tau* and *APOE*). It is likely that in the next several years, additional AD and FTD susceptibility genes will be identified. However, until most or all of these genes are identified, and their relative frequencies as a cause of AD or FTD in specific populations can be defined, genetic testing of at-risk family members can only be reasonably performed if there is a testable, clinically affected member available. If an affected pedigree member is available, their DNA can be screened for mutations in the known susceptibility genes. If mutations are found, this information can then be used for testing in at-risk family members. Where no mutations in the affected subject's DNA are found, it can be reasonably assumed that the disease is caused by mutations or polymorphisms in other AD/FTD genes yet to be identified, and mutation screening studies in at-risk family members would not be indicated. However, in the absence of a living affected member who can be tested, the screening for mutations in the DNA of at-risk family members is currently likely to be a fruitless task because, until all of the disease-causing genes for that trait are known, the failure to discover disease-related polymorphisms or mutations in the known genes can give no reassurance that the at-risk family member is not a carrier of a mutation in another susceptibility gene.

REFERENCES

Aerssens, J., Raeymaekers, P., Lilienfeld, S. *et al.* (2001). APOE genotype: no influence on galantamine treatment efficacy nor on rate of decline in Alzheimer's disease. *Dement Geriatr Cogn Disord*, **12**, 69–77.

Alberts, M. J., Graffagnino, C., McClenny, C., DeLong, D., Strittmatter, W., Saunders, A. M. and Roses, A. D. (1995). ApoE genotype and survival from intracerebral haemorrhage. *Lancet*, **346**, 575.

Arispe, N., Pollard, H. B. and Rojas, E. (1993). Giant multilevel cation channels formed by Alzheimer disease amyloid beta-protein A beta P-(1–40) in bilayer membranes. *Proc Natl Acad Sci USA*, **90**, 10573–7.

Bales, K. R., Verina, T., Dodel, R. C. *et al.* (1997). Lack of apolipoprotein E dramatically reduces amyloid beta-peptide deposition. *Nat Genet*, **17**, 263–4.

Bertram, L., Blacker, D., Mullin, K. *et al.* (2000). Evidence for genetic linkage of Alzheimer's disease to chromosome 10q. *Science*, **290**, 2302–3.

Bird, T. D., Lampe, T. H., Nemens, E. J. *et al.* (1988). Familial Alzheimer's disease in American descendants of the Volga Germans: probable genetic founder effect. *Ann Neurol*, **23**, 25–31.

Bird, T. D., Levy-Lahad, E., Poorkaj, P. *et al.* (1996). Wide range in age of onset for chromosome 1-related familial Alzheimer's disease. *Ann Neurol*, **40**, 932–6.

Bird, T. D., Sumi, S. M., Nemens, E. J. *et al.* (1989). Phenotypic heterogeneity in familial Alzheimer's disease: a study of 24 kindreds. *Ann Neurol*, **25**, 12–25.

Boulianne, G. L., Livne-Bar, I., Humphreys, J. M. *et al.* (1997). Cloning and characterization of the *Drosophila presenilin* homologue. *Neuroreport*, **8**, 1025–9.

Brown, J., Ashworth, A., Gydesen, S. *et al.* (1995). Familial non-specific dementia maps to chromosome 3. *Hum Mol Genet*, **4**, 1625–8.

Cao, X. and Sudhof, T. C. (2001). A transcriptionally correction of transcriptively active complex of APP

with Fe65 and histone acetyltransferase Tip60. *Science*, **293**, 115–20.

Checler, F. (2001). The multiple paradoxes of presenilins. *J Neurochem*, **76**, 1621–7.

Citron, M., Westaway, D., Xia, W. *et al.* (1997). Mutant presenilins of Alzheimer's disease increase production of 42-residue amyloid beta-protein in both transfected cells and transgenic mice. *Nat Med*, **3**, 67–72.

Clark, L. N., Poorkaj, P., Wszolek, Z. *et al.* (1998). Pathogenic implications of mutations in the tau gene in pallido-ponto-nigral degeneration and related neurodegenerative disorders linked to chromosome 17. *Proc Natl Acad Sci USA*, **95**, 13103–7.

Copley, T. T., Wiggins, S., Dufrasne, S. *et al.* (1995). Are we all of one mind? Clinicians' and patients' opinions regarding the development of a service protocol for predictive testing for Huntington disease. Canadian Collaborative Study for Predictive Testing for Huntington Disease. *Am J Med Genet*, **58**, 59–69.

Corder, E. H., Saunders, A. M., Risch, N. J. *et al.* (1994). Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. *Nat Genet*, **7**, 180–4.

Corder, E. H., Saunders, A. M., Strittmatter, W. J. *et al.* (1993). Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science*, **261**, 921–3.

Cupers, P., Orlans, I., Craessaerts, K., Annaert, W. and De Strooper, B. (2001). The amyloid precursor protein (APP)-cytoplasmic fragment generated by gamma-secretase is rapidly degraded but distributes partially in a nuclear fraction of neurones in culture. *J Neurochem*, **78**, 1168–78.

Davis, R. L., Shrimpton, A. E., Holohan, P. D. *et al.* (1999). Familial dementia caused by polymerization of mutant neuroserpin. *Nature*, **401**, 376–9.

De Strooper, B., Annaert, W., Cupers, P. *et al.* (1999). A presenilin-1-dependent gamma-secretase-like protease mediates release of Notch intracellular domain. *Nature*, **398**, 518–22.

De Strooper, B., Beullens, M., Contreras, B. *et al.* (1997). Phosphorylation, subcellular localization, and membrane orientation of the Alzheimer's disease-associated presenilins. *J Biol Chem*, **272**, 3590–8.

Duff, K., Eckman, C., Zehr, C. *et al.* (1996). Increased amyloid-beta42(43) in brains of mice expressing mutant presenilin 1. *Nature*, **383**, 710–3.

Ertekin-Taner, N., Graff-Radford, N., Younkin, L. H. *et al.* (2000). Linkage of plasma Abeta42 to a quantitative locus on chromosome 10 in late-onset Alzheimer's disease pedigrees. *Science*, **290**, 2303–4.

Farlow, M. R. (1997). Alzheimer's disease: clinical implications of the apolipoprotein E genotype. *Neurology*, **48**, S30–4.

Fassbender, K., Simons, M., Bergmann, C. *et al.* (2001). Simvastatin strongly reduces levels of Alzheimer's disease beta-amyloid peptides Abeta 42 and Abeta 40 in vitro and in vivo. *Proc Natl Acad Sci USA*, **98**, 5856–61.

Gotz, J., Chen, F., van Dorpe, J. and Nitsch, R. M. (2001). Formation of neurofibrillary tangles in *P301l tau* transgenic mice induced by Abeta 42 fibrils. *Science*, **293**, 1491–5.

Gu, Y., Chen, F., Sanjo, N. *et al.* (2003). APH-1 interacts with mature and immature forms of presenilins and nicastrin and may play a role in maturation of presenilin.nicastrin complexes. *J Biol Chem*, **278**, 7374–80.

Gydesen, S., Hagen, S., Klinken, L., Abelskov, J. and Sorensen, S. A. (1987). Neuropsychiatric studies in a family with presenile dementia different from Alzheimer and Pick disease. *Acta Psychiatr Scand*, **76**, 276–84.

Hendrie, H. C., Hall, K. S., Hui, S. *et al.* (1995). Apolipoprotein E genotypes and Alzheimer's disease in a community study of elderly African Americans. *Ann Neurol*, **37**, 118–20.

Herreman, A., Hartmann, D., Annaert, W. *et al.* (1999). Presenilin 2 deficiency causes a mild pulmonary phenotype and no changes in amyloid precursor protein processing but enhances the embryonic lethal phenotype of presenilin 1 deficiency. *Proc Natl Acad Sci USA*, **96**, 11872–7.

Hutton, M., Lendon, C. L., Rizzu, P. *et al.* (1998). Association of missense and 5′-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature*, **393**, 702–5.

Jarrett, J. T. and Lansbury, P. T., Jr. (1993). Seeding ''one-dimensional crystallization'' of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie? *Cell*, **73**, 1055–8.

Jick, H., Zornberg, G. L., Jick, S. S., Seshadri, S. and Drachman, D. A. (2000). Statins and the risk of dementia. *Lancet*, **356**, 1627–31.

Kamal, A., Almenar-Queralt, A., LeBlanc, J. F., Roberts, E. A. and Goldstein, L. S. (2001). Kinesin-mediated axonal transport of a membrane compartment containing beta-secretase and presenilin-1 requires APP. *Nature*, **414**, 643–8.

Kang, J., Lemaire, H.G., Unterbeck, A. *et al.* (1987). The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor. *Nature*, **325**, 733–6.

Katzman, R. and Kawas, C. (1994). In Terry, R.D., Katzman, R. and Bick, K.L. (eds.), *Alzheimer disease*, pp. 105–22. New York: Raven Press.

Kim, S.H., Wang, R., Gordon, D.J. *et al.* (1999). Furin mediates enhanced production of fibrillogenic ABri peptides in familial British dementia. *Nat Neurosci*, **2**, 984–8.

Kounnas, M.Z., Moir, R.D., Rebeck, G.W. *et al.* (1995). LDL receptor-related protein, a multifunctional ApoE receptor, binds secreted beta-amyloid precursor protein and mediates its degradation. *Cell*, **82**, 331–40.

Kwok, J.B., Taddei, K., Hallupp, M. *et al.* (1997). Two novel (*M233T* and *R278T*) presenilin-1 mutations in early-onset Alzheimer's disease pedigrees and preliminary evidence for association of presenilin-1 mutations with a novel phenotype. *Neuroreport*, **8**, 1537–42.

Lautenschlager, N.T., Cupples, L.A., Rao, V.S. *et al.* (1996). Risk of dementia among relatives of Alzheimer's disease patients in the MIRAGE study: What is in store for the oldest old? *Neurology*, **46**, 641–50.

Levitan, D. and Greenwald, I. (1995). Facilitation of lin-12-mediated signalling by sel–12, a *Caenorhabditis elegans S182* Alzheimer's disease gene. *Nature*, **377**, 351–4.

Lewis, J., Dickson, D.W., Lin, W.L. *et al.* (2001). Enhanced neurofibrillary degeneration in transgenic mice expressing mutant tau and APP. *Science*, **293**, 1487–91.

Lewis, J., McGowan, E., Rockwood, J. *et al.* (2000). Neurofibrillary tangles, amyotrophy and progressive motor disturbance in mice expressing mutant (P301L) tau protein. *Nat Genet*, **25**, 402–5.

Li, Y.J., Oliveira, S.A., Xu, P. *et al.* (2003). Glutathione S-transferase omega-1 modifies age-at-onset of Alzheimer disease and Parkinson disease. *Hum Mol Genet*, **12**, 3259–67.

Liddell, M.B., Lovestone, S. and Owen, M.J. (2001). Genetic risk of Alzheimer's disease: advising relatives. *Br J Psychiatry*, **178**, 7–11.

Lorenzo, A. and Yankner, B.A. (1994). Beta-amyloid neurotoxicity requires fibril formation and is inhibited by congo red. *Proc Natl Acad Sci USA*, **91**, 12243–7.

Lu, D.C., Rabizadeh, S., Chandra, S. *et al.* (2000). A second cytotoxic proteolytic peptide derived from amyloid beta-protein precursor. *Nat Med*, **6**, 397–404.

Lynch, T., Sano, M., Marder, K.S. *et al.* (1994). Clinical characteristics of a family with chromosome 17-linked disinhibition-dementia-Parkinsonism-amyotrophy complex. *Neurology*, **44**, 1878–84.

Maestre, G., Ottman, R., Stern, Y. *et al.* (1995). Apolipoprotein E and Alzheimer's disease: ethnic variation in genotypic risks. *Ann Neurol*, **37**, 254–9.

Mandelkow, E. and Mandelkow, E.M. (2002). Kinesin motors and disease. *Trends Cell Biol*, **12**, 585–91.

Martins, R.N., Turner, B.A., Carroll, R.T. *et al.* (1995). High levels of amyloid-beta protein from *S182 (Glu246)* familial Alzheimer's cells. *Neuroreport*, **7**, 217–20.

Mattson, M.P., Cheng, B., Davis, D. *et al.* (1992). beta-Amyloid peptides destabilize calcium homeostasis and render human cortical neurons vulnerable to excitotoxicity. *J Neurosci*, **12**, 376–89.

Mattson, M.P. and Goodman, Y. (1995). Different amyloidogenic peptides share a similar mechanism of neurotoxicity involving reactive oxygen species and calcium. *Brain Res*, **676**, 219–24.

Mayeux, R., Ottman, R., Maestre, G. *et al.* (1995). Synergistic effects of traumatic head injury and apolipoprotein-epsilon 4 in patients with Alzheimer's disease. *Neurology*, **45**, 555–7.

Mayeux, R., Saunders, A.M., Shea, S. *et al.* (1998). Utility of the apolipoprotein E genotype in the diagnosis of Alzheimer's disease. Alzheimer's Disease Centers Consortium on Apolipoprotein E and Alzheimer's Disease. *N Engl J Med*, **338**, 506–11.

McKeith, I.G., Burn, D.J., Ballard, C.G. *et al.* (2003). Dementia with Lewy bodies. *Semin Clin Neuropsychiatry*, **8**, 46–57.

Myers, A., Holmans, P., Marshall, H. *et al.* (2000). Susceptibility locus for Alzheimer's disease on chromosome 10. *Science*, **290**, 2304–5.

Newman, M.F., Croughwell, N.D., Blumenthal, J.A. *et al.* (1995). Predictors of cognitive decline after cardiac operation. *Ann Thorac Surg*, **59**, 1326–30.

Olichney, J.M., Hansen, L.A., Galasko, D. *et al.* (1996). The apolipoprotein E epsilon 4 allele is associated with increased neuritic plaques and cerebral amyloid angiopathy in Alzheimer's disease and Lewy body variant. *Neurology*, **47**, 190–6.

Olson, J.M., Goddard, K.A. and Dudek, D.M. (2002). A second locus for very-late-onset Alzheimer disease: a genome scan reveals linkage to 20p and epistasis between 20p and the amyloid precursor protein region. *Am J Hum Genet*, **71**, 154–61.

Perez-Tur, J., Froelich, S., Prihar, G. *et al.* (1995). A mutation in Alzheimer's disease destroying a splice

acceptor site in the *presenilin-1* gene. *Neuroreport*, **7**, 297–301.

Pericak-Vance, M. A., Bass, M. P., Yamaoka, L. H. *et al.* (1997). Complete genomic screen in late-onset familial Alzheimer disease. Evidence for a new locus on chromosome 12. *Jama*, **278**, 1237–41.

Pericak-Vance, M. A., Bebout, J. L., Gaskell, P. C., Jr. *et al.* (1991). Linkage studies in familial Alzheimer disease: evidence for chromosome 19 linkage. *Am J Hum Genet*, **48**, 1034–50.

Pericak-Vance, M. A., Grubber, J., Bailey, L. R. *et al.* (2000). Identification of novel genes in late-onset Alzheimer's disease. *Exp Gerontol*, **35**, 1343–52.

Pike, C. J., Burdick, D., Walencewicz, A. J., Glabe, C. G. and Cotman, C. W. (1993). Neurodegeneration induced by beta-amyloid peptides in vitro: the role of peptide assembly state. *J Neurosci*, **13**, 1676–87.

Podlisny, M. B., Citron, M., Amarante, P. *et al.* (1997). Presenilin proteins undergo heterogeneous endoproteolysis between Thr291 and Ala299 and occur as stable N- and C-terminal fragments in normal and Alzheimer brain tissue. *Neurobiol Dis*, **3**, 325–37.

Poorkaj, P., Bird, T. D., Wijsman, E. *et al.* (1998). Tau is a candidate gene for chromosome 17 frontotemporal dementia. *Ann Neurol*, **43**, 815–25.

Qiu, W. Q., Walsh, D. M., Ye, Z. *et al.* (1998). Insulin-degrading enzyme regulates extracellular levels of amyloid beta-protein by degradation. *J Biol Chem*, **273**, 32730–8.

Rebeck, G. W., Perls, T. T., West, H. L. *et al.* (1994). Reduced apolipoprotein epsilon 4 allele frequency in the oldest old Alzheimer's patients and cognitively normal individuals. *Neurology*, **44**, 1513–16.

Relkin, N. *et al.* (1996). Apolipoprotein E genotyping in Alzheimer's disease. National Institute on Aging/Alzheimer's Association Working Group. *Lancet*, **347**, 1091–5.

Rogaev, E. I., Sherrington, R., Rogaeva, E. A. *et al.* (1995). Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. *Nature*, **376**, 775–8.

Rogaeva, E., Premkumar, S., Song, Y. *et al.* (1998). Evidence for an Alzheimer disease susceptibility locus on chromosome 12 and for further locus heterogeneity. *Jama*, **280**, 614–18.

Rogaeva, E. A., Fafel, K. C., Song, Y. Q. *et al.* (2001). Screening for PS1 mutations in a referral-based series of AD cases: 21 novel mutations. *Neurology*, **57**, 621–5.

Roses, A. D. and Saunders, A. (1995). Head injury, amyloid beta and Alzheimer's disease. *Nat Med*, **1**, 603–4.

Runz, H., Rietdorf, J., Tomic, I. *et al.* (2002). Inhibition of intracellular cholesterol transport alters presenilin localization and amyloid precursor protein processing in neuronal cells. *J Neurosci*, **22**, 1679–89.

Sastre, M., Steiner, H., Fuchs, K. *et al.* (2001). Presenilin-dependent gamma-secretase processing of beta-amyloid precursor protein at a site corresponding to the S3 cleavage of Notch. *EMBO Rep*, **2**, 835–41.

Sato, S., Kamino, K., Miki, T. *et al.* (1998). Splicing mutation of presenilin-1 gene for early-onset familial Alzheimer's disease. *Hum Mutat*, Suppl 1, S91–4.

Saunders, A. M., Strittmatter, W. J., Schmechel, D. *et al.* (1993). Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. *Neurology*, **43**, 1467–72.

Scheuner, D., Eckman, C., Jensen, M. *et al.* (1996). Secreted amyloid beta-protein similar to that in the senile plaques of Alzheimer's disease is increased in vivo by the presenilin 1 and 2 and APP mutations linked to familial Alzheimer's disease. *Nat Med*, **2**, 864–70.

Schmechel, D. E., Saunders, A. M., Strittmatter, W. J. *et al.* (1993). Increased amyloid beta-peptide deposition in cerebral cortex as a consequence of apolipoprotein E genotype in late-onset Alzheimer disease. *Proc Natl Acad Sci USA*, **90**, 9649–53.

Scott, W. K., Hauser, E. R., Schmechel, D. E. *et al.* (2003). Ordered-subsets linkage analysis detects novel Alzheimer disease loci on chromosomes 2q34 and 15q22. *Am J Hum Genet*, **73**, 1041–51.

Selkoe, D. J. (1994). Normal and abnormal biology of the beta-amyloid precursor protein. *Annu Rev Neurosci*, **17**, 489–517.

Sherrington, R., Froelich, S., Sorbi, S. *et al.* (1996). Alzheimer's disease associated with mutations in presenilin 2 is rare and variably penetrant. *Hum Mol Genet*, **5**, 985–8.

Sherrington, R., Rogaev, E. I., Liang, Y. *et al.* (1995). Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature*, **375**, 754–60.

Sima, A. A., Defendini, R., Keohane, C. *et al.* (1996). The neuropathology of chromosome 17-linked dementia. *Ann Neurol*, **39**, 734–43.

Spillantini, M. G., Goedert, M., Crowther, R. A. *et al.* (1997). Familial multiple system tauopathy with presenile dementia: a disease with abundant neuronal and glial tau filaments. *Proc Natl Acad Sci USA*, **94**, 4113–18.

Spittaels, K., Van den Haute, C., Van Dorpe, J. *et al.* (1999). Prominent axonopathy in the brain and spinal cord of transgenic mice overexpressing four-repeat human tau protein. *Am J Pathol*, **155**, 2153–65.

St George-Hyslop, P., McLachlan, D.C., Tsuda, T. *et al.* (1994). Alzheimer's disease and possible gene interaction. *Science*, **263**, 537.

Strittmatter, W.J., Saunders, A.M., Schmechel, D. *et al.* (1993*a*). Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proc Natl Acad Sci USA*, **90**, 1977–81.

Strittmatter, W.J., Weisgraber, K.H., Huang, D.Y. *et al.* (1993*b*). Binding of human apolipoprotein E to synthetic amyloid beta peptide: isoform-specific effects and implications for late-onset Alzheimer disease. *Proc Natl Acad Sci USA*, **90**, 8098–102.

Terry, R.D., Masliah, E. and Hansen, L. (1994). In Terry, R.D., Masliah, E. and Hansen, L. (eds.), *Alzheimer disease*, pp. 179–96. New York: Raven Press.

Thinakaran, G., Borchelt, D.R., Lee, M.K. *et al.* (1996). Endoproteolysis of presenilin 1 and accumulation of processed derivatives in vivo. *Neuron*, **17**, 181–90.

Tierney, M.C., Szalai, J.P., Snow, W.G. *et al.* (1996). A prospective study of the clinical utility of ApoE genotype in the prediction of outcome in patients with memory impairment. *Neurology*, **46**, 149–54.

van Duijn, C.M., de Knijff, P., Cruts, M. *et al.* (1994). Apolipoprotein *E4* allele in a population-based study of early-onset Alzheimer's disease. *Nat Genet*, **7**, 74–8.

Vassar, R., Bennett, B.D., Babu-Khan, S. *et al.* (1999). Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science*, **286**, 735–41.

Vidal, R., Frangione, B., Rostagno, A. *et al.* (1999). A stop-codon mutation in the *BRI* gene associated with familial British dementia. *Nature*, **399**, 776–81.

Walter, J., Capell, A., Grunberg, J. *et al.* (1996). The Alzheimer's disease-associated presenilins are differentially phosphorylated proteins located predominantly within the endoplasmic reticulum. *Mol Med*, **2**, 673–91.

Wolfe, M.S. (2001). Presenilin and gamma-secretase: structure meets function. *J Neurochem*, **76**, 1615–20.

Wolfe, M.S., Xia, W., Ostaszewski, B.L. *et al.* (1999). Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and gamma-secretase activity. *Nature*, **398**, 513–17.

Wolozin, B., Kellman, W., Ruosseau, P., Celesia, G.G. and Siegel, G. (2000). Decreased prevalence of Alzheimer disease associated with 3-hydroxy-3-methyglutaryl coenzyme A reductase inhibitors. *Arch Neurol*, **57**, 1439–43.

Wong, P.C. (1996). Functions of the presenilins: generation and characterization of presenilin-1 null mice. *Society for Neuroscience*, **22**, 728.

Wszolek, Z.K., Pfeiffer, R.F., Bhatt, M.H. *et al.* (1992). Rapidly progressive autosomal dominant parkinsonism and dementia with pallido-ponto-nigral degeneration. *Ann Neurol*, **32**, 312–20.

Yankner, B.A., Duffy, L.K. and Kirschner, D.A. (1990). Neurotrophic and neurotoxic effects of amyloid beta protein: reversal by tachykinin neuropeptides. *Science*, **250**, 279–82.

Yasojima, K., Akiyama, H., McGeer, E.G. and McGeer, P.L. (2001). Reduced neprilysin in high plaque areas of Alzheimer brain: a possible relationship to deficient degradation of beta-amyloid peptide. *Neurosci Lett*, **297**, 97–100.

Ye, Y., Lukinova, N. and Fortini, M.E. (1999). Neurogenic phenotypes and altered Notch processing in *Drosophila Presenilin* mutants. *Nature*, **398**, 525–9.

Yu, G., Nishimura, M., Arawaka, S. *et al.* (2000). Nicastrin modulates presenilin-mediated Notch/glp-1 signal transduction and betaAPP processing. *Nature*, **407**, 48–54.

Zheng, H., Jiang, M., Trumbauer, M.E. *et al.* (1996). Mice deficient for the amyloid precursor protein gene. *Ann N Y Acad Sci*, **777**, 421–6.

30

Major psychiatric disorders in adult life

Amanda Elkin, Sridevi Kalidindi, Kopal Tandon and Peter McGuffin

Psychiatric disorders are common. The World Health Organization (WHO) estimates that at least one in four people will experience a clinically significant episode of psychiatric illness at some point in their lives. Although most such disorders are short lived and do not result in specialist care many cases become disabling and the WHO Global Burden of Disease Study (Murray and Lopez, 1997) has estimated that in health economic terms unipolar depression (UPD) vies with cardiovascular disease as the leading cause of disability in adults world wide. Schizophrenia and bipolar affective disorder (BPD) are also major public health problems which feature in the WHO's top ten of economically burdensome diseases. Between them, the affective disorders (UPD and BPD) and schizophrenia also account for over 60% of completed suicides. Therefore we will here focus on these three conditions as the main exemplars of common complex psychiatric disorders with substantial genetic contributions.

Clinical features and epidemiology

Unipolar depression (UPD) is so called because it consists of episodes of depressed mood whereas bipolar disorder (BPD) presents as episodes of both mania and depression. (A minority of patients with bipolar disorder have episodes of mania and no depressive episodes, but in terms of course, outcome, treatment response and pattern of illness in relatives they resemble typical bipolar disorder cases and are therefore classified as such rather than a ''unipolar mania''). In order to fulfill current criteria for depression (WHO, 1992; American Psychiatric Association, 1994) an episode must last at least two weeks but in practice most episodes last longer. In addition to low mood and low energy levels, other typical symptoms include anhedonia (the inability to enjoy), loss of self esteem, suicidal ideation, guilt and hopelessness. Biological symptoms are common such as disturbances of sleep, appetite, weight and sexual function. In mania the mood by contrast is one of elation (although irritability is also common). This is typically accompanied by racing thoughts, rapid speech, reckless overactivity, social and sexual disinhibition and overspending, sometimes to a ruinous level. There is almost always a reduced need for sleep and there are frequently expansive and grandiose ideas such as being in possession of special powers or abilities or of having become suddenly rich or famous.

UPD of sufficient severity to warrant hospital treatment has a lifetime risk in the UK of approximately 8% in women and 5% in men (McGuffin *et al.*, 1996) but epidemiological surveys in the United States suggest that lifetime prevalences in the community are much higher at around 30% in women and 10% in men (Kessler *et al.*, 2003). A female to male ratio of around two to three is a fairly constant finding except in those cases requiring in-patient treatment where the female preponderance is less (Sturt *et al.*, 1984). Classic BPD, also known as bipolar I, has a lifetime risk of less that 1% and a roughly equal sex incidence

but a less severe form, so called bipolar disorder II, affects up to 2% of the general population and there is a slight female preponderance.

Schizophrenia is characterized by hallucinations, delusions, disorder of the form of speech and psychomotor disturbance. These are collectively often called positive symptoms. Negative symptoms are less conspicuous but can be more disabling and include diminished motivation, social withdrawal and cognitive impairment, such that the old name for the syndrome was dementia praecox (early dementia). It typically arises in early adulthood with a lifetime risk of around 1% in both men and women, although the average age of onset is consistently earlier in men.

Genetic epidemiology

Affective disorders

There is no doubt that affective disorders tend to run in families. Among the first to remark on this was Kraepelin (1921), the founder of modern psychiatric classification, who noted that ''hereditary taint'' was apparent in 80% of his patients suffering from manic-depressive illness. The first studies to distinguish between UPD and BPD were comparatively recent. They were carried out by Angst (1966), a Swiss psychiatrist, and independently, in the same year, by Perris (1966) in Sweden. Perris found that there was a striking degree of homotypia, that is, a tendency for relatives of UP and BP probands to ''breed true'' and exhibit the same type of disorder. However, the findings of Angst were more complex. Among the relatives of probands with UPD there was an increase only of UPD, but among the relatives of BPD probands there was an excess of both UPD and BPD. Most subsequent studies have conformed to the pattern found by Angst (McGuffin and Katz, 1986).

It must be emphasized that, as with most common diseases, UPD and BPD do not show Mendelian patterns of transmission. Indeed even allowing for the possibility of incomplete penetrance the available data are incompatible with single locus inheritance (Craddock *et al.*, 1995). Also, familial aggregation does not necessarily mean that a disorder is genetically determined; relatives may resemble each other with respect to a behavioural trait or disorder because of shared genes or shared environment. Studies of monozygotic (MZ) and dizygotic (DZ) twins provide a means of assessing the relative contribution of each of these with a higher concordance rate in MZ than DZ twins being strongly suggestive of genetic effects. The results of two of the largest modern twin studies ascertained via hospital treated cases (Bertelsen *et al.*, 1977; McGuffin *et al.*, 1996; 2003*a*) are summarized in Table 30.1. Both show significantly higher MZ than DZ concordance. They also show that even in genetically identical individuals, MZ twins, homotypia was incomplete. The results therefore tend to support the impression from family studies that UPD and BPD differ quantitatively rather than being qualitatively distinct conditions.

Assuming that UPD and BPD are disorders to which the liability (Falconer, 1965) is contributed by multiple genetic and environmental factors, McGuffin *et al.* (1996) estimated that the heritability, or proportion of variance in liability to UPD attributable to additive genetic effects, was

Table 30.1. Probandwise concordance for unipolar (UP) and bipolar (BP) affective disorder in twins

Study		Proband	*n*	Co-twin UP %	Co-twin BP %	Co-twin Affective disorder %
Bertelson *et al.* (1977)	MZ	UP	35	43	9	64
		BP	34	18	62	79
	DZ	UP	17	18	6	24
		BP	37	11	8	19
McGuffin *et al.* (2003)	MZ	UP	68	44	2	46
		BP	30	27	40	67
	DZ	UP	109	20	0	20
		BP	37	14	5	19

48–75%. The heritability of BPD was somewhat higher at just over 80% (McGuffin *et al.*, 2003*a*). These authors then went on to perform a more complex statistical analysis, a bivariate form of structural equation modeling, to examine the relationship between UPD and BPD. They concluded that although there is substantial genetic overlap between the disorders, most of the liability to manic symptoms results from genes that are specific to BPD.

Focusing initially on women, Kendler and colleagues (Kendler *et al.*, 1992; 1993) studied a large population based sample that included 742 complete twin pairs (Kendler *et al.*, 1992). The heritability of major depression ranged from 21–45% with no evidence of shared (or familial) environmental effects. The higher heritability in the clinically based sample (McGuffin *et al.*, 1996) may reflect the possibility that twins ascertained through clinics are more severely ill than those sampled in the community. An indication that more severe depressive disorder is associated with higher heritability comes from a third twin series where the registry was based on former US military personnel from the Vietnam War era (Lyons *et al.*, 1998). It has also been suggested that, in addition to taking severity of depression into account, the age of onset might afford some clues as to the familiality. However, this is not supported by meta-analysis (Sullivan *et al.*, 2000). Another factor that needs to be considered in population based studies is failure to remember symptoms. Interestingly Kendler *et al.* (1993) found a large degree of diagnostic instability, or rather unreliability, in the Virginian female twins when their subjects were assessed at two time points. These authors therefore derived an ''index of caseness'' based on the number of depressive symptoms, treatment seeking, number of episodes, and degree of impairment. They found that there was better reliability and the heritability of lifetime major depression was greater for more restrictive definitions. The finding of multiple episodes being associated with higher heritability is a consistent result in other twin (McGuffin *et al.*, 1996; Lyons *et al.*, 1998) and family studies (Sullivan *et al.*, 2000). Kendler and colleagues (1993) then incorporated error of measurement into a structural equation model including both occasions of measurement. The estimated heritability of the liability to depression increased substantially to approximately 70%. More than half of what was estimated as environmental effects, when lifetime major depression was analysed on the basis of one assessment, was found to reflect measurement error when two assessments were used.

Adoption data are much less extensive but these also tend to support a genetic contribution to affective disorders. Mendlewicz and Rainer (1977) found that 28% of the biological parents of BPD adoptees had affective illness compared with 12% of their adopting parents. The rate of affective illness in the biological parents of adoptees did not differ significantly from that in the parents of BPD non-adoptees, of whom 26% were affected.

A smaller and less detailed study dealt mainly with UPD and concentrated on the offspring of patients with affective disorder (Cadoret, 1978). Adopted-away offspring of patients with affective illness had higher rates of affective disorder than those adoptees whose natural parents were psychiatrically well.

Based mainly on health insurance records, a Swedish adoption study carried out by Von Knorring *et al.* (1983) unexpectedly found little evidence of a genetic or family environmental component in affective illness. However, a more detailed study of Scandinavian data from Denmark (Wender *et al.*, 1986) found marked evidence of a genetic contribution to affective disorder. Assessments were made of hospital records of the biological relatives of adoptees with affective illness compared with the adoptive relatives and relatives of matched control adoptees. These were carried out blind to the status of the adoptee and showed an eightfold increase in UPD in the biological relatives of adoptees with affective illness compared with the other categories of relatives, as well as a 15-fold increase in the rate of suicide.

Table 30.2. Five recent twin studies of schizophrenia

Authors	Country	Ascertainment	Diagnostic criteria	MZ concordance (%)	DZ concordance (%)	Heritability %
Kläning (1996)	Denmark	Population register	ICD-10	7/16 (44)	2/19(11)	83
Cannon *et al.* (1998)	Finland	Population register	ICD-8/DSM IIIR	40/87(46)	18/195(9)	
Franzek and Beckmann (1998)	Germany	Hospital admissions	DSM IIIR	20/31(65)	7/25(28)	
Cardno *et al.* (1999*a*)	UK	Hospital register	DSM IIIR	20/47(43)	0/50(0)	84
			ICD-10	21/50(42)	1/58(2)	83
Tsujita *et al.* (1992)	Japan	Hospital admissions	DSM IIIR	11/22(50)	1/7(14)	
Combined			DSM IIIR	57/114(50)	4/97(4.1)	88
			ICD-10	28/66(42.4)	3/77(3.9)	83

Data from a review and meta-analysis by Cardno and Gottesman, 2000.

Schizophrenia

Schizophrenia has probably been more extensively investigated using family, twin and adoption studies than almost any other disorder. The family data are entirely consistent in showing a higher risk in relatives than in the general population. This averages around 10% in siblings and offspring but is usually somewhat lower in parents reflecting the fact that schizophrenia is associated with markedly reduced fecundity. Thus schizophrenia sufferers who have children tend to be those who had their offspring comparatively early and whose onset of disorder is comparatively late (Gottesman, 1991). As with the affective disorders, the pattern seen in families is not Mendelian and single gene explanations, even allowing for incomplete penetrance, are implausible mathematically (O'Rourke *et al.*, 1982) and statistically (McGue *et al.*, 1985).

There have now been many twin studies of schizophrenia all of which point to a definite genetic component. However, a criticism of studies carried out before the 1980s was that they used clinical diagnoses that were arguably more subjective and less reliable than modern, so called operational definitions of disorder such as those now contained in the American Psychiatric Association's Diagnostic and Statistical Manual (DSM) from the third edition (DSM III) onwards (American Psychiatric Association, 1980). To rectify this, Farmer *et al.* (1987) reassessed one of the best characterised schizophrenia twin series, that from the Maudsley Hospital twin register, London reported by Gottesman and Shields (1972). In fact applying the more rigorous DSM III definition led to a heritability estimate for liability to schizophrenia of around 80%, rather higher than had been found using a somewhat broader clinical diagnosis (McGue *et al.*, 1985). There have been five more recent twin studies where the cases have been defined using operational criteria (Table 30.2). One of these was a further updating and extension of the Maudsley series. Cardno *et al.* (1999) found proband-wise concordances of 20/47 (43%) in MZ twins compared with 0/50 (0%) in DZ twins using DSM IIIR (American Psychiatric Association, 1987) criteria. For the International Classification of Diseases, 10th edition (ICD 10) defined schizophrenia (World Health Organization, 1992), the rates were 21/50 (42%) in MZ twins and 1/58 (2%) in DZ twins.

Applying genetic model fitting, Cardno *et al.* (1999) estimated that the broad heritability, or proportion of variation in liability to schizophrenia resulting from genetic effects, was about 84% for the DSM IIIR and 83% for the ICD 10 definition of schizophrenia.

The adoption data on schizophrenia are particularly extensive and date from pioneering studies commenced in the 1960s by Heston (1966), Kety and others (1968). The time and energy devoted to adoption studies perhaps reflects the fact that schizophrenia was at the time seen as a battle ground both in the ''nature–nurture'' debate and in the debate about whether psychiatric disorders were ''real'' disorders at all. Essentially three designs have been employed in utilizing the 'natural experiment' of adoption. These are often referred to as the adoptees design, the adoptees' family design and the cross-fostering study. Results from all three designs are summarized in Table 30.3. For many, the adoption study results settled the debates about schizophrenia once and for all, leading Kety to famously remark that if ''schizophrenia is a myth, it is a myth with a strong genetic basis.''

Finding genes

The evidence from quantitative genetic approaches points consistently to a substantial genetic contribution to the major psychiatric disorders in adulthood. However the pattern of inheritance is complicated and is likely to involve multiple genes in an interplay with the environment. Therefore, next we need to address the question of whether such genes can be located and identified. The details of genetic mapping and positional cloning (Chapter 4) are essentially as follows. A chromosomal region containing a gene that confers susceptibility to a disorder is identified by linkage mapping or by linkage disequilibrium mapping. Using a variety of methods, the region is then narrowed down until the gene itself is identified. Subsequently, the mutations, or variations, that confer susceptibility to the disease are identified. Distribution, level of expression, and functions of the gene product can then be studied. The positional cloning approach is, in theory and increasingly in practice, straightforward in single gene disorders. The burning issue is whether the approach is feasible in common complex disorders that involve multiple genes.

Unfortunately, conventional linkage analysis requires several assumptions. These are that major gene effects (rather than just multiple small gene effects) exist, that there is some way of ensuring genetic homogeneity, and that the mode of transmission of the disorder is known. Concentrating on large multiply affected pedigrees with early onset has enabled these problems to be overcome in some complex disorders such as Alzheimer's disease and breast cancer, where single gene forms of these diseases have been identified, but to date this has not been the case with such psychiatric disorders as schizophrenia or BPD where the picture appears to be more complicated. An alternative to studying large families is sib-pair analysis, which has been successful in identifying susceptibility loci involved in such disorders as type 1 diabetes and is being used in a range of other common disorders. This type of study simply detects linkage by testing whether pairs of siblings, both affected by the disorder, share genetic marker alleles more than would be expected by chance. Focusing on affected sib pairs has several advantages, including not needing to make any assumptions about the mode of inheritance of the disease. The main drawback of sib pair analysis is that susceptibility genes of very small effect may require very large samples to be detected. For example, Williams *et al.* (1999) completed a whole genome scan involving nearly 200 sib pairs affected by schizophrenia. They found no definite positive linkage signals, but they were able to exclude a gene conferring a relative risk of two or more from over 80% of the genome. If it is correct that the genes involved in schizophrenia confer relative risks of less than two, then it may be that very large samples in the region of

Table 30.3. Principal findings from adoption studies: prevalence rates not age corrected

Study	Type of study	Diagnosis	Genetic relatives of a schizophrenic	Not genetically related to a schizophrenic
Heston (1966)	Adoptee	Schizophrenia	10.6% of 47 adoptees who had a schizophrenic biological mother	0% of 50 control adoptees
Rosenthal *et al.* (1968)	Adoptee	Schizophrenia spectrum disorder	18.8% of 69 children of schizophrenics raised by normals	10.1% of 79 control adoptees
Wender *et al.* (1974)	Cross-fostering	Schizophrenia spectrum disorder	18.8% of 69 children of schizophrenics raised by normals	10.7% of 28 children of controls raised by future schizophrenics
Kety (1983) Kety *et al.* (1994)	Adoptee's family: national sample (47 chronic schizophrenic adoptees)	Chronic and latent (DSM-II) schizophrenics	15.8% of 279 biological relatives of adopted-away schizophrenics	1.8% of 228 adoptive relatives of schizophrenics and relatives of control adoptees
Kendler and Gruenberg (1984, 1994)	Reassessment of Kety's data (31 adoptees with spectrum disorder)	DSM-III schizophrenia plus schizotypal personality disorder plus RDC schizoaffective disorder, mainly schizophrenic	14.4% of 209 biological relatives of adopted-away schizophrenics (23.5% in 1st-degree relatives, 9.9% in second-degree relatives)	3% of 299 adoptive relatives of schizophrenics and relatives of control adoptees
Tienari (1991)	Adoptee	Any form of psychosis	9% of 138 adoptees who had a schizophrenic biological parent	1.2% of 171 control adoptees

600–800 are required for linkage to be detected. Although such small effects of individual genes may seem at first sight surprising, it is not, in fact, out of keeping with the genetic epidemiology of schizophrenia (Gottesman, 1991). For example, the relative risk in a sibling of a schizophrenic is about ten, so it could be that if several additive genes are involved, none will individually have a relative risk of more than two. This is so because relative risks across different loci are multiplicative. Thus, for example, six loci each conferring a relative risk of 1.5 would give a total relative risk of over eleven.

The probability that genes of large effect are rare, perhaps non-existent, in the major psychiatric disorders and that liability generally results from the combined effects of multiple genes provides

the likely explanation for the fact that linkage studies have, until recently, provided a rather confusing and contradictory pattern. In an attempt to clarify the situation there have been attempts at meta-analyses. However, combining data from families selected in differing ways, from heterogeneous ethnic populations, using different markers for their genome scans and different methods of statistical analysis is problematic. Badner and Gershon (2002) adapted a method originally proposed by R. A. Fisher where *p*-values from multiple studies may be combined after correcting each value for the size of the region containing the minimum *p*-value. An alternative approach used by Levinson *et al.* (2003) divided the genome into segments or ''bins'' and used a ranking method to assess the degree of support across studies for linkage within each segment. The two methods as applied to schizophrenia data yielded overlapping but somewhat different results. The study of Badner and Gershon (2002) strongly supported the existence of susceptibility genes on chromosomes 8p, 13q and 22q while that of Levinson *et al.* (2003) and Lewis *et al.* (2003) favoured chromosomes 5q, 3p, 11q, 6p, 1q, 22q, 8p, 20q, and 14p. Thus the 8p and 22q regions in schizophrenia were supported by both meta-analyses but 8 other regions were supported by only one.

Badner and Gershon also applied their method to data on BPD and found that two of the ''schizophrenia'' regions were also strongly implicated here. That there might be some genes that contribute to the risk of both schizophrenia and BPD is at odds with traditional orthodoxy in both psychiatric classification and genetics. However it is in keeping with the results of one of the more recent twin analyses which found that, while it is highly probable that there are specific genes that contribute to the liability for schizophrenia and others that are specific to BPD, there is yet another set of genes that contributes to both syndromes (Cardno *et al.*, 2002). As we shall discuss later, there is evidence that at least one such schizophrenia/BPD gene has already been identified.

That said, the linkage data on BPD are by and large less extensive than those for schizophrenia. A meta-analysis using the method of Lewis *et al.* (2003) provided support for other BPD linkage regions on chromosomes 9p, 10q, 14q and 18 but this was suggestive rather than compelling. Genome wide systematic linkage scans in UPD are even fewer but are beginning to yield intriguing results. Some groups have studied a personality measure, neuroticism, which is strongly associated with the risk of clinical depression such that it is often difficult to tease the two apart in the context of family studies (Farmer *et al.*, 2002). A ''hotspot'' on chromosome 12 has been identified by three separate studies, two of UPD and the other of neuroticism. The study by Fullerton *et al.* (2003) which used siblings extremely discordant and concordant for neuroticism, found their strongest linkage with a marker on chromosome 12. Abkevich *et al.* (2003), studying Mormon families from Utah multiply affected by UPD, found highly significant evidence of linkage with a marker in the same vicinity. Oddly, most of the effect came from men with UPD so that these two studies must be seen as somewhat mutually supportive but as pointing to the need for a more intensive search in this part of the genome. A larger and more recent whole genome linkage scan, looking at 497 sib pairs concordant for recurrent major depression, supported evidence for linkage in this region. Importantly, this remained significant when the data was combined with previous reports (McGuffin *et al.*, 2005). Functionally, this region contains a gene, *DAO*, which has also been associated with schizophrenia and BPD. The same study also suggested evidence for linkage on chromosome 1p, and a gene associated with depression, *MTHFR*, is in the middle of this region. However, this region is different from that reported on the same chromosome by Fullerton *et al.* (2003) and Nurnberger *et al.* (2001). This latter study examined sibling pairs affected by either depression and/or alcoholism in families ascertained due to multiply affected members with alcoholism. Co-morbidity between depression

and alcoholism often occurs, although it is unclear as to what extent this is due to shared genetic factors. Winokur *et al.* (1978) proposed a category of ''Depression Spectrum Disease,'' which is characterised by a family history of alcoholism in first-degree relatives of those with a primary unipolar depressive disorder. It is of interest that Prescott *et al.* (2000), in their study of twins, found evidence for shared genetic factors influencing vulnerability to depression and alcoholism, although these factors appear to explain only a small proportion of the liability to each of the disorders.

More recently, Holmans *et al.* (2004) showed, for the first time, genomewide significant linkage to recurrent, early-onset major depression on chromosome 15q. This has, so far, been replicated by McGuffin *et al.* (2005) in another sib pair study but a putative susceptibility gene or genes in this region are currently unclear.

Positional candidate genes

Despite the fact that there remain ambiguities surrounding linkage findings in schizophrenia, several groups have been sufficiently emboldened to embark on detailed searches within putative linkage regions which has led to the publication within the past two years of a flurry of papers implicating several positional candidate genes in schizophrenia. Essentially, a positional candidate is a gene within a region of the genome identified by linkage studies that could plausibly be involved in the pathogenesis of the disorder. Having identified a candidate, its role is then examined further by testing for allelic association. This involves identifying DNA variants or polymorphisms within the gene and then testing whether a particular allele (an alternative variant) or a particular haplotype (a block of alleles carried together on the same chromosome) is more common in cases than in healthy controls. As an alternative to such case-control comparisons, a within-family design can be used. For example, affected individuals and their parents are genotyped and a test is performed as to whether there is differential transmission of a particular allele (or haplotype) from heterozygous parents to affected offspring (Sham and McGuffin, 2002).

As we have mentioned, several regions of the genome have been implicated by more than one linkage study and remain promising following meta-analyses.

The gene encoding dysbindin (*DTNBP1*), which maps to chromosome 6p22.3 is the most convincing susceptibility gene for schizophrenia to date. Since Straub *et al.* (2002) first identified this gene as involved in susceptibility a further 8 replication studies have been published. These included 14 independent samples, 11 of which showed positive associations. Data suggests that variation in *DTNBP1* expression confers susceptibility to schizophrenia via reduced expression (Bray *et al.*, 2005), thus suggesting a primary etiological mechanism for the disorder.

Neuregulin-1 (*NRG-1*), a gene which maps to chromosome 8p21, was first implicated in the susceptibility to schizophrenia by Stefansson *et al.* (2002). Two initial replications for the same haplotype were positive (Stefansson *et al.*, 2003 and Williams, N. *et al.*, 2003) whereas a further five studies provided support for the region but different haplotypes. More recently, there have also been three negative reports (Iwata *et al.*, 2002; Thistelton *et al.*, 2004; Hong *et al.*, 2004). Despite these, other evidence is in favour of *NRG1* having an etiological role including immunohistochemical findings (Law, 2003) human post-mortem brain studies (Davis *et al.*, 2003; Hakak *et al.*, 2001) and ''partial knockout'' mice heterozygous for *NRG-1* which show behavioral abnormalities consistent with schizophrenia. More work is necessary to identify the possible etiological significance of this region.

Association with markers around two overlapping genes on chromosome 13q, G72 (now known as DAO; D-amino-acid-oxidase) and G30, were found to be associated with schizophrenia by Chumanov *et al.* (2002). A number of replication

studies followed but there has not been any consensus regarding specific at-risk alleles or haplotypes. Furthermore, a large study comparing subjects with bipolar 1 disorder, schizophrenia and matched controls showed significant evidence for association of *DAO/G30* locus with BPD but not schizophrenia (Williams *et al.*, 2006).

Chromosome 22q11 is a region of interest, implicated in the etiology of schizophrenia both by linkage studies and by work on velocardiofacial syndrome (Murphy *et al.*, 1999). In this disorder, the genetic defect is a microdeletion of the 22q11 region, and sufferers have a 20–30% chance of developing a schizophrenia-like syndrome (Murphy, 2002). Thus, a number of candidate genes in this area have been identified. Firstly, Liu *et al.* (2002) identified a marker in the proline dehydrogenase, (*PRODH*) gene which showed a modest association with schizophrenia. Another group, Jacquet *et al.* (2002) supported this finding but a further study, with large sample size and high power to detect the association was negative (Williams, H. *et al.*, 2003).

A gene coding for catechol-O-methyltransferase (COMT) is also located on chromosome 22q11. The Val158Met polymorphism has been extensively studied, because of its key role in dopamine catabolism, and the results have been mixed. However, the largest and most recent meta-analysis (Munafo *et al.*, 2005) which looked at 18 studies, does not support the association with schizophrenia.

In addition, a sixth gene, regulator of G protein signalling 4 (*RGS4*), implicated by gene expression studies, is known to map to a linkage region on chromosome 1q and contains polymorphisms that have been found to be associated with schizophrenia (see reviews by Harrison and Owen, 2003; McGuffin *et al.*, 2003*b*; Elkin *et al.*, 2004; Horton *et al.*, 2006). Again, results are varied with a mixture of positive associations (Williams *et al.*, 2004; Chen *et al.*, 2004; Morris *et al.*, 2004) but also negative reports (Sobell *et al.*, 2005; Cordeiro *et al.*, 2005).

Also, on chromosome 1q is another region of interest, not only to schizophrenia but also to BPD and UPD. The original group that suggested this (Blackwood *et al.*, 2001) discovered an extended Scottish family with a balanced chromosomal translocation (1;11) significantly linked to a clinical phenotype that included schizophrenia and affective disorders. The genes, *DISC 1* and *DISC 2*, which are directly disrupted by the breakpoint on chromosome 1 were postulated to have a role in the development of these disease phenotypes. Since then a number of studies (Hogkinson *et al.*, 2004; Callicot *et al.*, 2005; Zhang *et al.*, 2005; Kockelkorn *et al.*, 2004) have found evidence for association for these disorders, although for different markers or haplotypes. More work is necessary.

Functional candidate genes

Lest they be forgotten in the recent wave of interest in positional candidates, two genes involved in dopamine and serotonin pathways but not implicated by linkage have much longer histories and have been extensively investigated. These are the serotonin receptor 5HT2A and the dopamine receptor DRD3, which is one of the dopamine D2 class of receptors. They have long been regarded as prime candidates in schizophrenia because of what is known about the mode of action of antipsychotic drugs. The older generation of so called typical antipsychotics (e.g. haloperidol, chlorpromazine) have the common property of blocking dopamine D2 receptors in the brain. The more recently introduced atypical antipsychotics (e.g. clozapine, olanzapine) block both D2 type and 5HT2A receptors.

5HT2A has a single nucleotide polymorphism in the first exon termed T102C. This does not result in an amino acid change, but it may result in lower expression of the *C* allele in the temporal cortex of normal individuals and schizophrenia patients (Polesskaya and Sakalov, 2002), suggesting that a functional effect is conferred by this polymorphism or another variant in linkage disequilibrium. Several, but not all, case-control studies of patients

with schizophrenia that investigated this polymorphism, have found an over-representation of the *C* allele in cases compared to controls. A meta-analysis combining over 3000 subjects showed a positive association (Williams *et al.*, 1997) which was of small effect, but highly significant. Since this, there have been further negative studies, but another recent meta-analysis of 28 5HT2A studies (Lohmueller *et al.*, 2003) lends further weight to this finding being a true one.

The gene encoding the dopamine D3 receptor (*DRD3*) has a *Ser9Gly* polymorphism in exon 1. Homozygosity for this polymorphism was found to show an association with schizophrenia in a combined study in France and in Wales (Stefansson *et al.*, 2002). Subsequent studies have yielded results that lend support to this finding but several others have been negative. However the most recent meta-analysis (Steffanson *et al.*, 2002), which includes 48 studies of *DRD3* containing over 5000 individuals, gives credence to the association being true with a small but significant effect.

Two other functional candidate genes have also received recent interest, including several replication attempts for each. They are *V-AKT* murine thymoma viral oncogene homolog 1 (*AKT-1*) and glutamate receptor, metabotropic 3 (*GRM3*), both located in the 22q11 region. The evidence for the former as a susceptibility gene for schizophrenia is encouraging (Ohtusuki *et al.*, 2004; Ikeda *et al.*, 2004; Schwab *et al.*, 2005) and further investigation is certainly warranted. The evidence for the latter includes several reports of association (Marti *et al.*, 2002; Fujii *et al.*, 2003; Egan *et al.*, 2004; Chen *et al.*, 2005; Norton *et al.*, 2005), but as yet, no replication of a marker or haplotype, thus further studies are needed

Conclusions

There is abundant evidence from family, twin and adoption studies that genes make an important contribution to schizophrenia and to bipolar and unipolar affective disorders. However the patterns of transmission are complicated and single gene forms are rare, or non-existent. There is also a substantial environmental component in all three disorders but this seems to consist mainly of non-shared (non-familial) environmental effects. Recent statistical analyses using structural equation modelling have challenged traditional orthodoxies about etiological distinctions, particularly between schizophrenia and BPD. It now appears that although there are genes that are specific to each of these syndromes there are also overlapping genes that contribute to both. Recent molecular genetic studies suggest that such genes have already been identified; *DAO/G30* and *DISC1/DISC 2*.

Molecular genetic studies of schizophrenia and the affective disorders have proven to be difficult but are now beginning to yield consistent results, particularly in schizophrenia. The findings so far regarding the role of NRG-1 and dysbindin in schizophrenia are the most consistent. Two genes within the linkage/VCFS microdeletion region of chromosome 22q, COMT and PRODH2, have also been the subject of much interest. Although the data relating to the former was initially very encouraging a recent meta-analysis does not support this association. The data on PRODH2 are also somewhat confusing at present with several positive pointers but one fairly large negative study. Putative functional candidate genes within this 22q11 region also include V-AKT murine thymoma viral oncogene homolog 1 (*AKT-1*) and glutamate receptor, metabotropic 3 (GRM3). Thus, despite some contradictory results, progress in elucidating the molecular genetic basis of schizophrenia is beginning to make clear strides.

These recent findings take us into a new and potentially exciting phase of ''post genomic psychiatry'' presenting several big challenges. The first is to find out the functional significance of variations in the susceptibility genes discovered so far. Reviews by Harrison and Owen (2003) and by Manji *et al.* (2003) offer possible explanations and provocative speculations on the role of the

''new genes'' in neurotransmission and signal transduction. Harrison and Owen point out that dysbindin, neuregulin, DAA/G30 and RGS4 are all likely to have modulating effects on glutamatergic transmission while Manji *et al.* offer a unifying theory that takes into account dopamine and glutamate signalling pathways particularly those involving N-methyl-D-aspartate (NMDA) receptors.

A second challenge is to discover what other susceptibility genes there are, for there must be more to come, and to discover how gene coactions and interactions occur at a molecular level. One promising newcomer is *PPP3CC*, a gene which maps (like neuregulin) to chromosome 8p21 and which codes for CaNAγ, a catalytic subunit of calcineurin. Manji *et al.* (2003) point out that this could regulate both dopaminergic and glutamatergic dysfunction and abnormalities of neuronal development that are postulated in schizophrenia. *PPP3CC* has been shown to be associated with schizophrenia in two samples (Gerber *et al.*, 2003) but further association study results are awaited. A third challenge is a re-evaluation of the whole edifice of current psychiatric classification. We predict that changes will be required in the light of genetic evidence. For example there is already a strong suggestion from analysis of twin data (Cardno *et al.*, 2002) that although there are genetic effects that are specific to schizophrenia and bipolar disorder, there are also overlapping genes that contribute to both syndromes. The results from linkage (Badner and Gershon, 2002) and from positional candidate genes studies (Hattori *et al.*, 2003) are beginning to bear this out.

REFERENCES

Abkevich, V., Camp, N.J., Hensel, C.H. *et al.* (2003). Predisposition locus for major depression at chromosome 12q22–12q23.2. *American Journal of Human Genetics*, **73**, 1271–81.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders, Third edition (DSM III)*. Washington, DC: American Psychiatric Association.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders, Third edition, revised (DSM IIIR)*. Washington, DC: American Psychiatric Association.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental Disorders, Fourth Edition (DSM IV)*. Washington, DC: American Psychiatric Association.

Angst, J. (1966). Zur aetilogie und nosologie endogener depressiver psychosen. *Monographen ans der Neurologie und Psychiatrie*, N112. Berlin: Springer-Verlag.

Badner, J.A. and Gershon, E.S. (2002). Meta-analysis of whole-genome linkage scans of bipolar disorder and schizophrenia. *Molecular Psychiatry*, **7**, 405–11.

Bertelson, A., Havald, B. and Hauge, M. (1977). A Danish twin study of manic depressive disorder. *British Journal of Psychiatry*, **130**, 330–51.

Blackwood, D., Fordyce, A., Walker, M. *et al.* (2001). Schizophrenia and affective disorders – cosegregation with a translocation at chromosome 1q42 that directly disrupts brain-expressed genes: clinical and P300 findings in a family. *American Journal of Medical Genetics*, **69**, 428–33.

Bray, N., Preece, A., Williams, N. *et al.* (2005). Haplotypes at the dystrobrevin binding protein 1 (*DTNBP1*) gene locus mediate risk for schizophrenia through reduced *DTNBP1* expression. *Human Molecular Genetics*, **14**, 1947–54.

Cadoret, R. (1978). Evidence for genetic inheritance of primary affective disorder in adoptees. *American Journal of Psychiatry*, **133**, 463–6.

Callicott, J., Straub, R., Pezawas, L. *et al.* (2005). Variation in DISC1 affects hippocamal structure and function and increases risks for schizophrenia. *Proceedings of the National Academy of Science USA*, **102**, 8627–32.

Cannon, T.D., Kaprio, J., Lonnqvist, J., Huttunen, M. and Koskenvuo, M. (1998). The genetic epidemiology of schizophrenia in a Finnish twin cohort: a population-based modeling study. *Arch Gen Psychiatry*, **55**, 67–74.

Cardno, A.G., Gottesman, II (2000). Twin studies of schizophrenia: from bow-and-arrow concordances to Star Wars Mx and functional genomics. *American Journal of Medical Genetics*, **97**, 12–17.

Cardno, A.G., Marshall, E.J., Coid, B. *et al.* (1999). Heritability estimates for psychotic disorders: The Maudsley Twin Psychosis series. *Archives of General Psychiatry*, **56**, 162–8.

Cardno, A. G., Rijsdijk, F. V., Sham, P. C., Murray, R. M. and McGuffin, P. (2002). A twin study of genetic relationships between psychotic symptoms. *American Journal of Psychiatry*, **159**, 539–45.

Chen, Q., He, G., Chen, O. *et al.* (2005). A case-control study of the relationship between the metabotropic glutamate receptor 3 gene and schizophrenia in the Chinese population. *Schizophrenia Research*, **73**, 21–6.

Chen, X., Dunham, C., Kendler, S. *et al.* (2004). Regulator of G-protein signalling for (*RGS4*) gene is associated with schizophrenia in Irish high density families. *American Journal of Medical Genetics B Neuropsychiatric Genetics*, **129**, 23–6.

Chumankov, I., Blumefield, M., Guerassimenko, O. *et al.* (2002). Genetic and physiological data implicating the new human gene G72 and the gene for D-amino acid oxidase in schizophrenia. *Proceedings of the National Academy of Science USA*, **99**, 13675–80.

Cordeiro, Q., Talkowski, M., Chowdari, K. *et al.* (2005). Association and linkage analysis of *RGS4* polymorphisms with schizophrenia and bipolar disorder in Brazil. *Genes Brain and Behaviour*, **4**, 45–50.

Craddock, N., Khodel, V., van Eerdewegh, P. and Reich, T. (1995). Mathematical limits of multilocus models: the genetic transmission of bipolar disorder. *American Journal of Human Genetics*, **57**, 690–702.

Davis, K., Stewart, D., Friedman, J. White matter changes in schizophrenia. (2003). *Archives in General Psychiatry*, **60**, 443–56.

Elkin, A., Kalidindi, S. and McGuffin, P. (2004). Have schizophrenia genes been found? *Current Psychiatric Opinion*, **17**, 107–13.

Egan, M. F., Straub, R., Goldberg, T. *et al.* (2004). Variation in the GRM3 affects cognition, prefrontal glutamate, and risk for schizophrenia. *Proceedings of the National Academy of Science USA*, **101**, 12604–9.

Falconer, D. S. (1965). The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Annals of Human Genetics*, **29**, 51–76.

Farmer, A. E., McGuffin, P. and Gottesman, I. I. (1987). Twin concordance for DSM-III schizophrenia: Scrutinizing the validity of the definition. *Archives of General Psychiatry*, **44**, 634–41.

Farmer, A., Redman, K., Harris, T. *et al.* (2002). Neuroticism, extraversion, life events and depression. The Cardiff Depression Study. *British Journal of Psychiatry*, **181**, 118–22.

Franzek, E. and Beckmann, H. (1998). Different genetic background of schizophrenia spectrum psychoses: a twin study. *American Journal of Psychiatry*, **155**, 76–83.

Fujii, Y., Shibata, H., Kikuta, R. *et al.* (2003). Positive association of polymorphisms in the metabotropic glutamate receptor type 3 gene (GRM3) with schizophrenia. *Psychiatric Genetics*, **13**, 71–6.

Fullerton, J., Cubin, M., Tiwari, H. *et al.* (2003). Linkage analysis of extremely discordant and concordant sibling pairs identifies quantitative-trait loci that influence variation in the human personality trait neuroticism. *American Journal of Human Genetics*, **72**, 879–90.

Gerber, D. H., Hall, D., Miyakawa, T. *et al.* (2003). Evidence for association of schizophrenia with genetic variation in the 8p21.3 gene, PPP3CC, encoding the calcineurin gamma subunit. *Procedures of the National Academy of Science USA*, **100**, 8933–8.

Gottesman, I. I. and Shields, J. (1972). *Schizophrenia and genetics. A twin study vantage point.* New York and London: Academic Press.

Gottesman, I. I. (1991). *Schizophrenia genesis.* New York: W. H. Freeman.

Hakak, Y., Walker, J. and Li, C. (2001). Genome-wide expression analysis reveals dysregulation of myelantion-related genes in chronic schizophrenia. *Proceedings of the National Academy of Science USA*, **98**, 4746–51.

Harrison, P. J. and Owen, M. J. (2003). Genes for schizophrenia? Recent findings and their pathophysiological implications. *Lancet*, **361**, 417–19.

Hattori, E., Liu, C., Badner, J. A. *et al.* (2003). Polymorphisms at the *G72/G30* gene locus, on 13q33, are associated with bipolar disorder in two independent pedigree series. *American Journal of Human Genetics*, **72**, 1131–40.

Heston, L. L. (1966). Psychiatric disorders in foster home reared children of schizophrenic mothers. *British Journal of Psychiatry*, **112**, 819–25.

Hodgkinson, C., Goldman, D., Jaeger, J. *et al.* (2004). Disrupted in schizophrenia 1 (DISC1): association with schizophrenia, schizoaffective disorder, and bipolar disorder. *American Journal of Human Genetics*, **75**, 862–72.

Holmans, P., Zubenko, G., Crowe, R. *et al.* (2004). Genomewide significant linkage to recurrent, early-onset major depressive disorder on chromosome 15q. *American Journal of Human Genetics*, **74**, 1154–67.

Horton, N., Williams, H. and Owen, M. J. (2006). An update on the genetics of schizophrenia. *Current Opinion in Psychiatry*, **19**, 158–64.

Hong, C., Huo, S., Liao, D. *et al.* (2004). Case-control and family-based association studies between the neuregulin 1 (*Arg38Gln*) polymorphism and schizophrenia. *Neuroscience Letters*, **366**, 158–61.

Ikeda, M., Iwata, N., Suzuki, T. *et al.* (2004). Association of AKT1 with schizophrenia confirmed in a Japanese population. *Biological Psychiatry*, **56**, 698–700.

Iwata, N., Suzuki, T., Ikeda, M. *et al.* (2004). No association with the neuregulin 1 haplotype to Japanese schizophrenia. *Molecular Psychiatry*, **9**, 126–7.

Jaquet, H., Raux, G., Thibaut, F. *et al.* (2002). PRODH mutations and hyperprolinaemia in a subset of schizophrenia patients. *Human Molecular Genetics*, **11**, 2243–9.

Kendler, K. S. and Gruenberg, A. M. (1984). An independent analysis of the Danish Adoption Study of Schizophrenia. VI. The relationship between psychiatric disorders as defined by DSM-III in the relatives and adoptees. *Archives of General Psychiatry*, **41**, 555–64.

Kendler, K. S., Gruenberg, A. M. and Kinney, D. K. (1994). Independent diagnoses of adoptees and relatives as defined by DSM-III in the provincial and national samples of the Danish Adoption Study of Schizophrenia. *Archives of General Psychiatry*, **51**, 456–68.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C. and Eaves, L. J. (1992). A population based twin study of major depression in women: the impact of varying definitions of illness. *Archives of General Psychiatry*, **49**, 257–66.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C. and Eaves, L. J. (1993). The lifetime history of major depression in women: reliability of diagnosis and heritability. *Archives of General Psychiatry*, **50**, 863–70.

Kessler, R. C., Berglund, P., Demler, O. *et al.* (2003). National Comorbidity Survey Replication. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *Journal of the American Medical Association*, **289**, 3095–105.

Kety, S. S. (1983). Mental illness in the biological and adoptive relatives of schizophrenic adoptees: findings relevant to genetic and environmental factors in etiology. *American Journal of Psychiatry*, **140**, 720–7.

Kety, S., Rosenthal, D., Wender, P. H. and Shulsinger, F. (1968). The type and prevalence of mental illness in the biological and adopted families of adopted schizophrenics. *Journal of Psychiatry Research*, **6**, 345–62.

Kety, S. S., Wender, P. H., Jacobsen, B. *et al.* (1994). Mental illness in the biological and adoptive relatives of schizophrenic adoptees. Replication of the Copenhagen Study in the rest of Denmark. *Archives of General Psychiatry*, **51**, 442–55.

Kockelkorn, T., Arai, M., Matsumoto, H. *et al.* (2004). Association study of polymorphisms in the 5′ upstream region of human *DISC1* gene with schizophrenia. *Neuroscience Letters*, **368**, 41–5.

Kläning, U. (1996). *Schizophrenia in twins: incidence and risk factors*. Unpublished Doctoral dissertation, University of Aarhus, Denmark.

Kraepelin, E. (1921). *Manic-depressive insanity and paranoia*. Edinburgh: E. & S. Livingstone.

Law, A. (2003). Neuregulin-1 in the human brain. *American Journal of Psychiatry*, **160**, 1392.

Levinson, D. F., Levinson, M. D., Segurado, R. and Lewis, C. M. (2003). Genome scan meta-analysis of schizophrenia and bipolar disorder. Part I: Methods and power analysis. *American Journal of Human Genetics*, **73**, 17–33.

Lewis, C. M., Levinson, D. F., Wise, L. H. *et al.* (2003). Genome scan meta-analysis of schizophrenia and bipolar disorder. Part II: Schizophrenia. *American Journal of Human Genetics*, **73**, 34–48.

Liu, H., Heath, S., Sobin, C. *et al.* (2002). Genetic variation at the 22q11 PRODH2/DGCR6 locus presents an unusual pattern and increases susceptibility to schizophrenia. *Proceedings of the National Academy of Sciences*, **99**, 3717–22.

Lohmueller, K., Pearce, C. and Pike, M. (2003). Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nature Genetics*, **33**, 177–82.

Lyons, M. J., Eisen, S. A., Goldberg, J. *et al.* (1998). A registry-based twin study of depression in men. *Archives of General Psychiatry*, **55**, 468–72.

Manji, H. K., Gottesman, I. I. and Gould, T. D. (2003). Signal transduction and genes-to-behaviors pathways in psychiatric diseases. *Science STKE*, **207**, 49.

Marti, S., Cichon, S., Propping, P. and Nothen, M. (2002). Metabotropic glutamate receptor 3 (*GRM3*) gene variation is not associated with schizophrenia or bipolar affective disorder in the German population. *American Journal of Medical Genetics*, **114**, 46–50.

McGue, M., Gottesman, I.I. and Rao, D.C. (1985). Resolving genetic models for the transmission of schizophrenia. *Genetic Epidemiology*, **2**, 99–110.

McGuffin, P. and Katz, R. (1986). Nature, nurture, and affective disorder. In Deakin, J.F.W. ed., *The biology of affective disorders*, pp. 26–52. London: Gaskell Press, Royal College of Psychiatrists.

McGuffin, P., Katz, R., Watkins, S. and Rutherford, J. (1996). A hospital based twin register of the Heritability of DSM-IV Unipolar Depression. *Archives of General Psychiatry*, **53**, 129–36.

McGuffin, P., Knight, J., Breen, G. *et al.* (2005). Whole genome linkage scan of recurrent depressive disorder from the depression network study. *Human Molecular Genetics*, **14**, 3337–45.

McGuffin, P., Rijsdijk, F., Andrew, M. *et al.* (2003*a*). The heritability of bipolar affective disorder and the genetic relationship to unipolar depression. *Archives of General Psychiatry*, **60**, 497–502.

McGuffin, P., Tandon, K. and Corsico, A. (2003*b*). Linkage and association studies of schizophrenia. *Current Psychiatry Reports*, **5**, 121–7.

Mendlewicz, J. and Rainer, J.D. (1977). Adoption study supporting genetic transmission in manic depressive illness. *Nature*, **268**, 327–9.

Morris, D., Rodgers, A., McGhee, K. *et al.* (2004). Confirming *RGS4* as a susceptibility gene for schizophrenia. *American Journal of Medical Genetics B Neuropsychiatric Genetics*, **125**, 50–3.

Munafo, M., Bowes, L., Clark, T. and Flint, J. (2005). Lack of association of the COMT (*VAL158/108 Met*) gene and schizophrenia: a meta-analysis of case-control studies. *Molecular Psychiatry*, **10**, 765–70.

Murphy, K., Jones, L. and Owen, M. (1999). High rates of schizophrenia in adults with velo-cardio facial syndrome. *Archives of General Psychiatry*, **56**, 940–5.

Murphy, K. (2002). Schizophrenia and velo-cardio facial syndrome. *Lancet*, **359**, 426–30.

Murray, C.J. and Lopez, A.D. (1997). Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. *Lancet*, **349**, 1436–42.

Norton, N., Williams, H., Dwyer, S. *et al.* (2005). No evidence for the association between polymorphisms in GRM3 and schizophrenia. *BMC Psychiatry*, **5**, 23.

Nurnberger, J.I. Jr., Foroud, T., Flury, L. *et al.* (2001). Evidence for a locus on chromosome 1 that influences vulnerability to alcoholism and affective disorder. *American Journal of Psychiatry*, **158**, 718–24.

Ohtsuki, T., Inada, T. and Arinami, T. (2004). Failure to confirm association between *AKT1* haplotype and schizophrenia in a Japanese case-control population. *Molecular Psychiatry*, **9**, 981–3.

O'Rourke, D.H., Gottesman, I.I., Suarez, B.K., Rice, J. and Reich, T. (1982). Refutation of the general single-locus model for the etiology of schizophrenia. *American Journal of Human Genetics*, **34**, 630–49.

Perris, C. and D'Elia, G. (1966). A study of bipolar (manic-depressive) and unipolar recurrent depressive psychoses X: mortality, suicide, and life cycles. *Acta Psychiatrica Scandinavica*, **42** (194 suppl), 172–83.

Polesskaya, O. and Sakalov, B. (2002). Differential expression of the ''C'' and ''T'' alleles of the *5HT2A* receptor gene in the temporal cortex of normal individuals and schizophrenics. *Journal of Neuroscience Research*, **67**, 812–22.

Prescott, C.A., Aggen, S.H. and Kendler, K.S. (2000). Sex-specific genetic influences on the comorbidity of alcoholism and major depression in a population-based sample of US twins. *Archives of General Psychiatry*, **57**, 803–11.

Rosenthal, D., Wender, P.H., Kety, S.S. *et al.* (1968). Schizophrenics' offspring reared in adoptive homes. In Rosenthal, D. and Kety, S.S. (eds.), *The transmission of schizophrenia*, Oxford: Pergamon Press.

Schwab, S., Hoefgen, B., Hanses, C. *et al.* (2005). Further evidence for association of variance in the *AKT1* gene with schizophrenia in a sample of European sib-pair families. *Biological Psychiatry*, **58**, 446–50.

Sham, P. and McGuffin, P. (2002). Linkage and association. In McGuffin, P., Owen, M.J. and Gottesman, I.I. (eds.), *Psychiatric genetics and genomics*, pp. 55–73. Oxford: Oxford University Press.

Sobell, J., Richard, C., Wirshing, D. and Heston, L. (2005). Failure to confirm association between RGS4 haplotypes and schizophrenia in Caucasians. *American Journal of Medical Genetics B Neuropsychiatric Genetics*, **139**, 23–7.

Stefansson, H., Sigurdsson, E. and Steinthorsdottir, V. (2002). Neuregulin 1 and susceptibility to schizophrenia. *American Journal of Human Genetics*, **71**, 877–92.

Stefansson, H., Sarginson, J., Kong, A. *et al.* (2003). Association of neuregulin 1 with schizophrenia confirmed in a Scottish population. *American Journal of Human Genetics*, **72**, 83–7.

Straub, R., Jiang, Y., MacLean, C. *et al.* (2002). Genetic variation in the 6p22.3 gene *DTNBP1*, the human ortholog of the mouse dysbindin gene, is associated

with schizophrenia. *American Journal of Human Genetics*, **71**, 337–48.

Sturt, E., Kumakura, N. and Der, G. (1984). How depressing life is–life-long morbidity risk for depressive disorder in the general population. *Journal of Affective Disorders*, **7**, 109–22.

Sullivan, P. F., Neal, M. C. and Kendler, K. S. (2000). Genetic epidemiology of major depression: review and meta-analysis. *American Journal of Psychiatry*, **157**, 1552–62.

Thisleton, D., Webb, B., Neale, M. C., *et al.* (2004). No evidence for linkage or association of neuregulin-1 (*NRG*-1) with disease in the Irish study of high-density schizophrenia families (ASHDSF). *Molecular Psychiatry*, **9**, 777–83.

Tienari, P. (1991). Interaction between genetic vulnerability and family environment: the Finnish adoptive family study of schizophrenia. *Acta Psychiatrica Scandinavica*, **84**, 460–5.

Tsujita, T., Okazaki, Y., Fujimaru, K. *et al.* (1992). Twin concordance rate of DSM-III-R schizophrenia in a new Japanese sample. Abstracts, Seventh International Congress on Twin Studies (152). Tokyo, Japan.

von Knorring, A. L., Cloninger, C. R., Bohman, M. and Sigvardsson, S. (1983). An adoption study of depressive disorders and substance abuse. *Archives of General Psychiatry*, **40**, 943–50.

Wender, P. H., Rosenthal, D., Kety, S. S., Schulsinger, F. and Welner, J. (1974). Crossfostering. A research strategy for clarifying the role of genetic and experiential factors in the etiology of schizophrenia. *Archives of General Psychiatry*, **30**, 121–8.

Wender, P. H., Kety, S. S., Rosenthal, D. *et al.* (1986). Psychiatric disorders in the biological and adoptive families of adopted individuals with affective disorders. *Archives of General Psychiatry*, **43**, 923–9.

Williams, H., Williams, N., Spurlock, G. *et al.* (2003). Association between PRODH and schizophrenia is not confirmed. *Molecular Psychiatry*, **8**, 644–5.

Williams, J., McGuffin, P. and Nothen, M. (1997). Meta-analysis of the association between the 5-HT2a receptor T102C polymorphism and schizophrenia. *Lancet*, **348**, 1221.

Williams, N., Green, E., Macgregor, S. *et al.* (2006). *Archives of General Psychiatry*, **63**, 366–73.

Williams, N., Preece, A., Spurlock, G. *et al.* (2003). Support for genetic variation in neuregulin 1 and susceptibility to schizophrenia. *Molecular Psychiatry*, **8**, 485–7.

Williams, N., Preece, A., Spurlock, G. *et al.* (2004). Support of RGS4 as a susceptibility gene for schizophrenia. *Biological Psychiatry*, **55**, 192–5.

Williams, N. M., Rees, M. I., Holmans, P. *et al.* (1999). A two-stage genome scan for schizophrenia susceptibility genes in 196 affected sibling pairs. *Human Molecular Genetics*, **8**, 1729–39.

Winokur, G., Behar, D., Vanvalkenburg, C. and Lowry, M. (1978). Is a familial definition of depression both feasible and valid? *Journal of Nervous and Mental Disease*, **166**, 764–8.

World Health Organization. (1992). *International statistical classification of diseases and related health problems, 1989 Revision*. 10th Edition (ICD 10). Geneva: World Health Organization.

Zhang, X., Tochigi, M., Ohashi, J. *et al.* (2005). Association Study of the DISC1/TRAX locus with schizophrenia in a Japanese population. *Schizophrenia Research*, **79**, 175–80.

31

Speech and language disorders

Gabrielle Barnby and Anthony J. Monaco

Introduction

Linguistic impairment is a core diagnostic criterion for a number of developmental disorders, such as autism and specific language impairment (SLI). Uncovering the genetic mechanisms responsible for susceptibility to language impairments will be essential to our understanding of the central deficit underlying speech and language disorders. An overlap between autistic and SLI phenotypes has been proposed based on the shared characteristic of pragmatic language impairment (PLI) in some cases. PLI describes inappropriate communication within a social context and can be observed in a subset of individuals diagnosed with both autism and SLI (Bishop and Norbury, 2002). The existence of a phenotypic overlap suggests that a shared genetic susceptibility may be responsible for some aspects of language delay. This is supported by a higher rate of autism in siblings of probands with SLI than in the general population (3%:0.17% respectively) (Tomblin *et al.*, 2003). It has also been reported that siblings of autistic individuals have higher than expected language and communication deficits (Folstein *et al.*, 1999). Whether these observations indicate a common genetic pathway or intermediate phenotypes common to both disorders is unclear. For example, Bishop and Norbury (2002) found a group of children with PLI also met criteria for autism, whereas another group with PLI including stereotyped language and abnormal intonation were otherwise social and communicative.

The approach taken to identify genes underlying speech and language disorders depends on the genetic model indicated by segregation analysis. If a speech and language phenotype is inherited according to a straightforward Mendelian pattern then parametric linkage analysis followed by haplotype analysis and positional cloning can be performed. This methodology was used to successfully identify a critical linkage region (SPCH1) and subsequently the disease gene (*FOXP2*) involved in the rare case of an autosomal dominant monogenic speech and language disorder. More commonly, no clear mode of inheritance is indicated and a multifactorial and polygenic aetiology is implied. In the latter case, different strategies are necessary to locate susceptibility genes; such as non-parametric linkage analysis and association studies. This chapter will describe the clinical characteristics and epidemiology of a severe monogenic form of speech and language disorder, and the more complex disorders of SLI and autism. Evidence will be presented to demonstrate the involvement of a genetic component in each disorder. Molecular genetic approaches to identifying the genes giving rise to susceptibility in speech and language disorders will also be described including genome screen data, candidate gene studies and, where possible, proposed molecular mechanisms. Examining the evidence from these studies will help address the hypothesis that there are commonalities in the genetic pathways of speech and language impairments across a range of disorders.

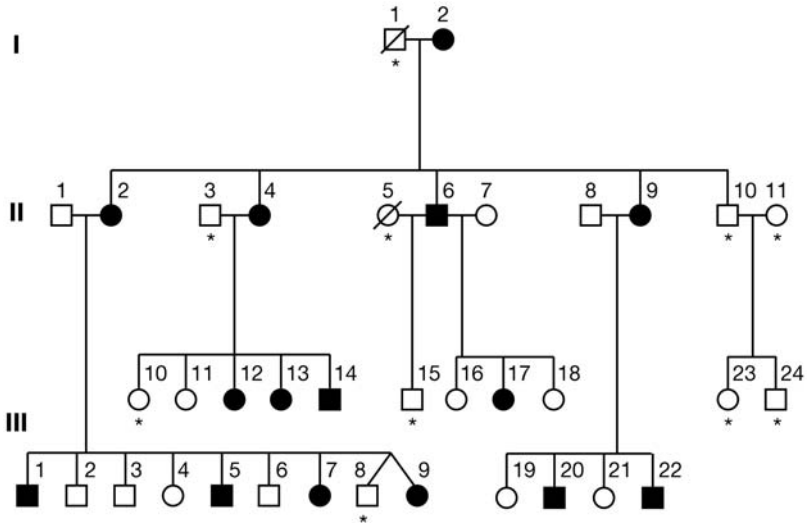

Figure 31.1 Pedigree of the KE family. Affected individuals are shaded, Asterisks indicate DNA unavailable for analysis. (From Fisher *et al.* (1998) © Nature America Inc.)

Epidemiology of speech and language disorders

A rare monogenic speech and language disorder

The KE family is a large three-generation pedigree segregating a severe form of speech and language disorder. Approximately half of the members of the KE family are affected, including males and females, and an autosomal dominant transmission is indicated (Hurst *et al.*, 1990), as shown in Figure 31.1. The KE family is unique because the speech and language disorder segregates in the family as a single Mendelian trait. This presented an opportunity to localize and identify a gene affecting the development of speech and language using traditional positional cloning.

KE family phenotype

There is an ongoing debate over the core primary defect in the KE family because affected family members appear to have deficits in a wide range of language and grammar abilities (Vargha-Khadem *et al.*, 1995). Affected family members have severe orofacial dyspraxia, which impedes co-ordination of facial muscles during speech and non-speech movements. Alongside these co-ordination impairments, grammatical ability is also markedly reduced and a number of family members, with and without speech and language impairment, possess IQs below the normal range. Despite the complex phenotype presented by the KE family, the striking pattern of Mendelian inheritance indicates that a single gene is responsible.

Patient CS

An unrelated patient, CS, has also been described who displays a similar severe orofacial dyspraxia and possesses impairments in expressive and receptive language ability. The phenotype of patient CS closely resembles that seen in affected members of the KE family, although there is no evidence of intellectual impairment.

Specific language impairment (SLI)

SLI is estimated to affect ~2–7% of pre-primary school children and is diagnosed if significant language deficits are present despite adequate educational opportunity and normal nonverbal intelligence (Law *et al.*, 1998). Alternative diagnoses of medical and developmental disorders such as deafness and autism must be excluded before a diagnosis of SLI is given. Features such as the orofacial dyspraxia found in the KE family would exclude a diagnosis of SLI. A continuum of language impairment exists, depending on the level of articulation and verbal expression, the ability to produce speech sounds, and verbal comprehension of others. SLI has been classified into subtypes, but it is possible that the variability seen in the phenotype reflects severity of impairment and the developmental stage of the child rather than an expression of specific deficits. Early evidence from a longitudinal study of language-impaired twins found that the majority (65.9%) of individuals diagnosed with early language delay at two years had resolved spontaneously by four years. This indicates that the provision of speech and language treatment at a very early age may be inefficient and unnecessary (Dale *et al.*, 2003) and that differentiating between delay and impairment, especially in young children, is a complex task. It remains unresolved how best to identify individuals requiring intervention where there is concern for speech and language delay. Some children will benefit from speech therapy and compensate for their difficulties as adults, but many do not and experience lifelong language difficulties and associated social problems.

Evidence for a genetic component to SLI

There is an increased frequency of language impairment in first-degree relatives of probands with SLI and the rate of SLI in monozygotic twins is elevated compared to dizygotic twins. Table 31.1 shows SLI twin concordance and prevalence rates. Family studies of SLI do not observe a traditional Mendelian pattern of inheritance or detect simple co-segregation between genotype and phenotype indicating that multiple genes may be involved. The heritability of SLI was estimated to be 0.45–0.95 by early studies (Bishop *et al.*, 1995; Tomblin and Buckwalter, 1998) but a high heritability of language impairment has not been consistently found. However, recent analysis of data from a large sample of 6-year-old twins finds a high heritability which supports the earlier findings (Bishop *et al.*, 2006). Interestingly, male twin pairs have a higher concordance rate than female twin pairs and SLI is increased in males compared to females, with a ratio of ~2:1, indicating a possible sex-linked effect (or sex-limited susceptibility). The substantial variation in prevalence rates reported for SLI are to a large extent due to study design and clinical assessment. As well as twin concordance studies, evidence for a genetic component to SLI also comes from estimating the heritabilities of quantitative measures of language ability, such as the Clinical Evaluation of Language Fundamentals-Revised (CELF-R) (Bishop *et al.*, 1999) and the MacArthur Communicative Development Inventory (MCDI) (Dale *et al.*, 1998). Additionally, prospective studies of infants with a family history of language impairment have found an association between temporal processing ability and later language outcome, indicating the presence of genetic risk before expressive language acquisition (Benasich and Spitz, 1998).

Table 31.1. SLI and autism twin concordance and prevalence rates

	SLI	Autism
Monozygotic twins	90–100%	60% (92%)
Dizygotic twins	50–70%	0% (10%)
Siblings/first degree relatives	18–42%	3.94% (5–6%)
Population prevalence	2–7%	0.13% (0.63%)

Figures in brackets are for a broader spectrum of cognitive and social abnormalities characteristic of autism.
(Lamb *et al.*, 2002; SLIC, 2002; Fombonne, 2005).

Clinical measures for SLI

Measures commonly used to assess language ability include the CELF-R, which consists of tests for expressive language and receptive syntactic-language abilities (the ordering of words in phrases and sentences). An additional test of non-word repetition, where subjects repeat nonsense words of increasing complexity such as ''contra-mptionist,'' has also been developed. This test is a good index of phonological short-term memory and thought to be a sensitive measure for language impairment (Bishop *et al.*, 1999). The Test of Language Development (TLD), which assesses comprehension, expression, syntax, grammar and phonology, can also be used to measure language impairment. Alternatively, the Children's Communication checklist (CCC) allows parents and teachers to evaluate communication impairments and avoid lengthy clinical assessments of children.

Summary

Epidemiological studies of SLI indicate the presence of a strong genetic component. However, a multigenic rather than simple genetic mode of inheritance is suggested and a highly variable phenotype is observed. This precludes the use of simple parametric analysis based on a known model of inheritance. The problems of clearly distinguishing between affected and unaffected family members adds further complexity to dissecting the genetics of SLI and indicates that a quantitative trait approach may be most powerful for determining the location of susceptibility genes.

Autism spectrum disorders

Core autism is composed of a triad of behavioral abnormalities including impairments in verbal and non-verbal communication, restricted social interaction and repetitive or stereotyped behaviours. Onset is around two years of age and clinical diagnosis can usually be made by four years of age. Similarly to SLI there are sex differences in prevalence, with three to four times as many males affected compared to females (Chakrabarti and Fombonne, 2001). The autism spectrum encompasses a range of disability outside the core definition of the disorder; the different diagnoses within this spectrum are collectively known as pervasive developmental disorders (PDDs). The population prevalence of autism is approximately 13 per 10 000 and this rises to almost 60 per 10 000 when all PDDs are considered (Fombonne, 2005). Affected individuals require different levels of assistance, from full time care in the most severely affected individuals, to independent living in high functioning individuals with Asperger's syndrome.

Evidence for a genetic component to autism

Support for a genetic contribution to autism comes from both twin and family studies. Monozygotic twins have a higher concordance rate compared to dizygotic twins and siblings of a proband with autism are over 20 times more likely to have autism compared with the general population, as shown in Table 31.1. The considerable decline in prevalence rates with decreased degrees of relatedness indicates a polygenic mode of inheritance and latent class analysis suggests between three and ten genes are likely to be involved (Pickles *et al.*, 1995).

Clinical measures for autism

The principal clinical diagnostic instruments for autism are the Autism Diagnostic Observation Schedule–Generic (ADOS-G), an interactive interview for observing social and communicative behaviour, and the Autism Diagnostic Interview-Revised (ADI-R), an interview for caregivers of individuals with autism. Traits examined include: frequency of eye contact; quality of imaginative play; ability to form friendships; and rigidity in routines. Assessment for autism also generally includes standardised measures of both performance and verbal IQ. These instruments are able

to distinguish between core autism and other PDDs. The presence of a broader autistic phenotype in family members has also been reported with mild abnormalities present in one area of functioning, such as language (Pickles *et al.*, 2000). It is worth noting that formally measuring language levels in autistic individuals is challenging because of both their social deficits and variations in their ability to perform in structured tasks (Tager-Flusberg, 2000).

Further observations on autism

Almost 30% of individuals with autism never develop functional language (Rapin, 1998) and additional individuals lose acquired language skills following developmental regression around two years of age (Lainhart *et al.*, 2002). The phenomenon of autistic regression, although it accounts for only 20% of cases, has led to speculation on possible environmental triggers for autism, including vaccine administration; particularly the MMR (measles, mumps, rubella) vaccination. There is no evidence from large-scale epidemiological studies that the MMR is associated with autism. In fact one of the largest studies to date, in a Danish cohort of over half a million children, found that receiving the vaccination was mildly protective for autism (Madsen *et al.*, 2002). For a review of MMR and autism epidemiological studies see (Fombonne and Cook, 2003; Rutter, 2005).

Clues to the neurological basis of autism have emerged from neuropathological examinations, although there is no consensus between studies. Forebrain abnormalities, cerebellar pathogenesis and cortical dysgenesis have all been described by different studies but none of these findings have been replicated (Lord and Bailey, 2002). However, a consistent observation in samples of autistic individuals is macrocephaly, and there is good evidence that autistic individuals follow an abnormal pattern of brain growth (Aylward *et al.*, 2002; Hazlett *et al.*, 2005). Neurochemical differences have also been described for autistic individuals but there is little understanding of any clear causal link with autism.

Summary

Autism is one of the most heritable developmental disorders, estimates of broad heritability are over 0.90 (Bailey *et al.*, 1995), with multiple interacting genes likely to be involved in its pathogenesis. As with SLI, a broad spectrum of disability exists and there are difficulties in delineating clear diagnostic categories. Because of the low frequency of autism, establishing the heritability of particular measures for quantitative analysis is problematic. However, employing strict diagnostic criteria makes possible a non-parametric qualitative approach to linkage analysis.

Molecular genetic approaches to speech and language disorders

SPCH1

Identification of the SPCH1 *gene – FOXP2*

The *SPCH1* locus on chromosome 7q31 was identified in 1998 following a genome-wide search for linkage in the KE family. A maximum pairwise log of odds (LOD) score of 6.22 was observed at the microsatellite marker D7S486. Genotyping additional markers and analysis of haplotypes for recombination events in affected individuals identified a 5.6cM critical region, containing the disease locus (*SPCH1*) (Fisher *et al.*, 1998). A positional cloning strategy was taken and in silico analysis identified 20 known genes and a further 50 anonymous transcripts within the region. Determining which of these transcripts was the *SPCH1* gene was aided by the precise localization of the translocation breakpoint of patient CS, as shown in Figure 31.2. Patient CS carries a de novo balanced reciprocal translocation t(5;7) (q22;q31.2) in the *SPCH1* region (Lai *et al.*, 2000). Detailed molecular analysis using fluorescence in situ hybridization (FISH) identified

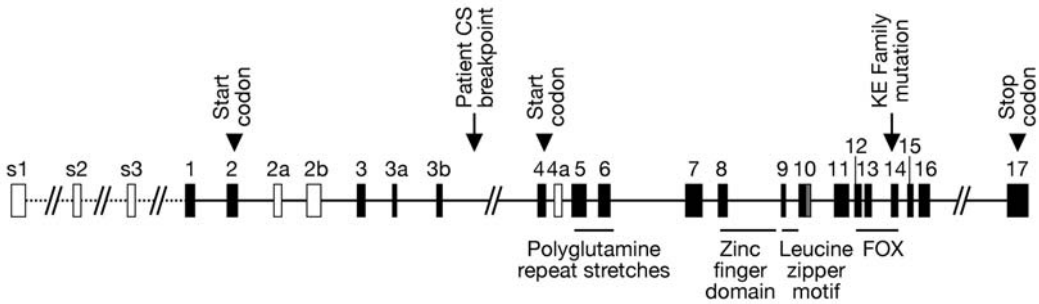
Start codon
Patient CS breakpoint
Start codon
KE Family mutation
Stop codon
s1 s2 s3 1 2 2a 2b 3 3a 3b 4 4a 5 6 7 8 9 10 11 12 13 14 15 16 17
Polyglutamine repeat stretches
Zinc finger domain
Leucine zipper motif
FOX

Figure 31.2 Genomic structure and protein domains of *FOXP2*. Introns >50 kb are indicated by //.

a single *BAC* clone that gave signals from both derivative chromosomes indicating it spanned the 7q31 breakpoint and may contain the disrupted *SPCH1* gene. The gene disrupted by the translocation in CS, within the *BAC* clone, was characterized and designated *FOXP2*. *FOXP2* was subsequently sequenced in KE family members and a G/A nucleotide change causing an arginine to histidine (R553H) substitution was found in exon 14 (Lai *et al.*, 2001). The variant co-segregated perfectly with speech and language disorder and was not found in 364 control chromosomes.

FOXP2 genomic structure and function

FOXP2 is a member of the FOX family of transcription factors, which share a characteristic DNA-binding domain composed of a forked/winged-helix (FOX). These transcription factors are involved in cellular differentiation and proliferation, signal transduction and pattern formation. *FOXP2* also contains a long stretch of glutamine residues; expansion of polyglutamine domains have previously been implicated in other neurological diseases (Cummings and Zoghbi, 2000). However, no significant variation in the *FOXP2* polyglutamine stretch has been detected in normal individuals or samples of individuals affected with neurodegenerative diseases, SLI and autism (Bruce and Margolis, 2002; Newbury *et al.*, 2002). The polyglutamine stretch is probably stabilized by the mixture of CAG and CAA codons it contains.

The KE family mutation (*R553H*) occurs in the most conserved region of the forkhead domain and is adjacent to the residue that makes direct contact with target DNA promoter sequences. The CS translocation is localized in intron three of the gene. Initially 17 exons with alternatively spliced exon 3a and 3b were described from a 6.5 kb transcript. Three additional 5′ exons, that may be alternatively spliced for exon 1, have subsequently been identified along with three internal exons 2a, 2b and 4a and a longer transcript of exon 10 (Bruce and Margolis, 2002). The principal protein isoform of *FOXP2* is 715 amino acids long and includes two polyglutamine stretches, a zinc finger motif and a forkhead domain; alternative splicing results in isoforms with extra amino acids (Bruce and Margolis, 2002). *FOXP2* is believed to span a genomic region of at least 603 kb and further exons and splice variants may yet be identified, Figure 31.2 shows the genomic structure of *FOXP2*.

Expression studies of FOXP2

Previous studies of FOX family members have found that they are important regulators of embryogenesis. Northern blot analysis of *FOXP2* shows widespread expression in adult and fetal tissues including strong expression in fetal brain and lung (Lai *et al.*, 2001). The expression of *FOXP2* and the mouse homolog *Foxp2* has been studied in the developing brain of mouse and human.

The expression patterns were similar in mouse and human and suggest that *FOXP2* has a conserved role in the development of corticostriatal and olivocerebellar circuits involved in motor control. Expression is found in the basal ganglia, which has been hypothesized to contribute to motor sequencing and procedural memory-like skills (Ferland *et al.*, 2003; Lai *et al.*, 2003; Lai *et al.*, 2003; Takahashi *et al.*, 2003). *FOXP2* expression was also detected in the cerebral cortex, which is interesting given the importance of the insular cortex for speech co-ordination. The spatial expression of *FOXP2* was found to correspond with structural abnormalities detected by magnetic resonance imaging (MRI) and functional differences detected by positron emission tomography (PET) in brains of affected members of the KE family. *FOXP2* is expressed in the developing caudate nucleus within the basal ganglia where a bilateral reduction of gray matter is found in affected KE family members (Vargha-Khadem *et al.*, 1998). The relative importance of the expression of *FOXP2* in the developing brain has yet to be determined.

Proposed molecular mechanisms

Reports on other members of the FOX family, which cause autosomal dominant disease, suggest that mutations disrupting the function of the forkhead domain lead to haplo-insufficiency or unbalanced gene dosage during embryonic development. For example, mutations in *FOXC1* cause hereditary glaucoma and can involve either deletion or duplication of the *FOX* gene (Nishimura *et al.*, 2001). The molecular mechanisms of how disruption of the DNA binding site in *FOXP2* affects neural development are yet to be described. Additional expression studies and the development of animal models will begin to uncover the role of *FOXP2* in the pathways leading towards the uniquely human traits of speech and language (Marcus and Fisher, 2003). In addition, investigation of the intraspecific variation of the human *FOXP2* gene showed that it contains amino-acid coding changes and a pattern of nucleotide polymorphism, which indicate that *FOXP2* has been the target of selection during recent human evolution, consistent with the emergence of spoken language (Enard *et al.*, 2002).

SLI and autism

Although large extended pedigrees affected by autism and SLI do exist, no clear Mendelian pattern of inheritance has been consistently observed. There are also very few families affected by autism over several generations because autistic individuals very rarely have children. The strategy for locating genomic regions likely to harbor susceptibility genes has relied upon whole genome screens for linkage. The majority of these screens used non-parametric or model free statistics to analyse linkage data, because the genetic model for SLI or autism remains unknown, but various parametric models have also been used for analysis. A larger number of studies have been carried out for autism compared to SLI but for both disorders the identification of susceptibility genes remains elusive and narrowing critical regions has also proved challenging.

Genome screens for SLI

Three whole genome screens have been performed for SLI and several susceptibility loci have been identified (Bartlett *et al.*, 2002; SLIC, 2002; SLIC, 2004). The study by Bartlett *et al.* included branches of five large families with a history of language and reading impairments. Following assessment with a wide battery of tests, individuals were assigned an affection status based on their spoken language, reading ability or history of language impairment. Multiple parametric analyses were performed using different penetrance levels and disease allele frequencies, which corresponded to a population prevalence of ~7%. Significant evidence for linkage was found at 13q21 for reading impairment with a two-point LOD of

3.62 (under a recessive model). There was also evidence for linkage at 2p22 for language impairment and 17q23 for reading impairment with two-point LODs of 2.28 and 1.92 (under recessive and dominant models respectively). Bartlett *et al.* have extended their initial study to examine regions of linkage previously found for both SLI and autism in new samples ascertained for SLI (Bartlett *et al.*, 2004). They suggest that studies of individuals with SLI may help to identify a genetic component which is common to both disorders, given that some children with autism have language abilities similar to children with SLI. They find evidence for linkage at 13q21 and more modest linkage signals on 7q and 2q which overlap with autism genome screen results in their extended sample.

In contrast to Bartlett *et al.*, the SLI Consortium (SLIC) followed a non-parametric quantitative trait locus (QTL) linkage approach where no prior assumptions were made about the genetic basis of the disease. It is important that phenotypes with detectable heritabilities are selected for QTL analysis. The SLIC sample included 98 families containing a proband with standard language scores ≥1.5 SD below the mean for their age. Three quantitative language measures were analysed: the CELF-R receptive and expressive language scales and a non-word repetition test. Two regions with maximum multipoint LOD scores of 3.55 were detected on chromosome 16q and 19q. The first locus on 16q was linked to the non-word repetition measure while the second locus on 19q was linked to the CELF-R expressive language score. A second genome screen using an independent sample of 86 multiplex families by the SLI Consortium supports these findings with a MLS of 2.84 generated on 16q and a MLS of 2.31 on 19q (SLIC, 2004). Combining the data from both SLI Consortium genome screens provides strong evidence for a susceptibility gene on 16q with an MLS of 7.46. The positions of the loci implicated in SLI by these genome screens are shown in Figure 31.3.

Genome screens for autism

Nine whole genome screens have been performed for autism to date and loci have been implicated on all chromosomes except 11, 12, 14, 20, and 21(MLS >1.0). The lack of suitable quantitative measures for autism has resulted in the use of large numbers of affected sib pairs for qualitative linkage analysis and the majority of studies have employed non-parametric linkage analysis. Figure 31.3 shows linkage reported by whole genome screens for autism. The top four linkages reported to date are on 3q25–27 with a maximum multipoint LOD score (MMLS) of 4.21, 2q21–33 with a MMLS score of 3.74, 7q22.2–31 with a MMLS of 3.20 and 13q22 with a MMLS of 3.0 (Barrett *et al.*, 1999; IMGSAC, 2001*a*; Auranen *et al.*, 2002). The most consistently reported linkage from genome screens is to chromosome 7q near the *SPCH1* locus, as shown in Figure 31.4. Convincing replication of linkage results for complex genetic disorders has been equivocal and attempts have been made to perform a meta-analysis of published linkage results. These meta-analyses have used methods such as: ranking linkage results across genome screens to analyse between scan heterogeneity, analysis allowing for the posterior probability of linkage in different subsets, and using multiple scan probabilities. (Badner and Gershon, 2002; Bartlett *et al.*, 2005; Cantor *et al.*, 2005; Trikalinos *et al.*, 2006). These meta-analyses of data find support for linkage reported to chromosome 1, 7q and 17q11 but suffer from a number of statistical limitations and none include all published genome screens. Additional analysis of autism linkage data has focused on sex-limited and parent-of-origin effects by looking at linkage signals driven by same-sex sibling pairs and parental sharing (Lamb *et al.*, 2005).

Several groups have analyzed genome screen data based on phenotypic subsets with the aim of delineating more homogeneous groups and strengthening linkage signals. Measures relating to language ability in particular have been used to subgroup families. Two studies have grouped sib

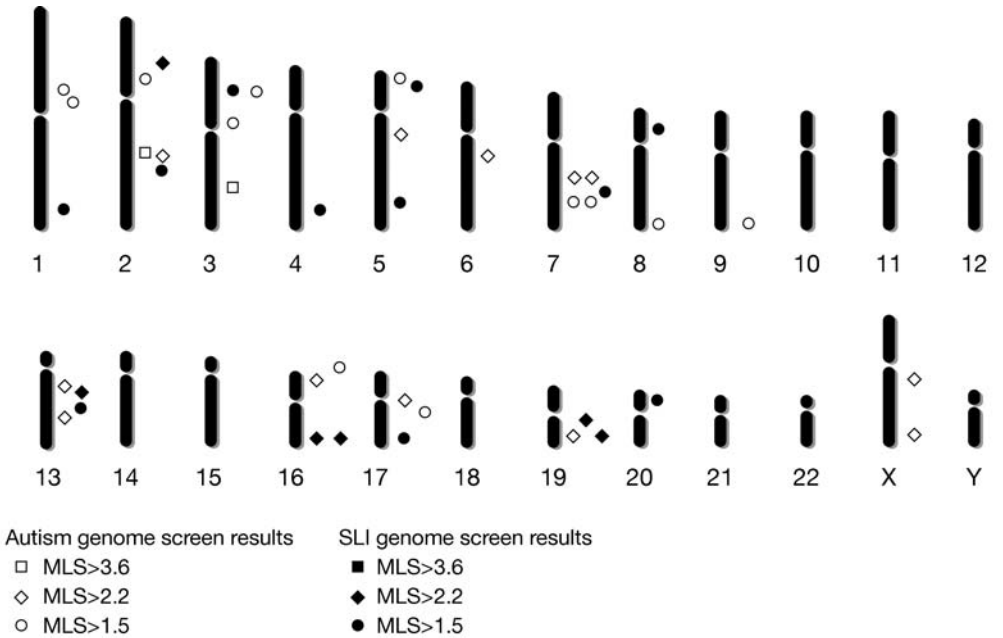

Figure 31.3 Loci detected by SLI and autism genome screens.

pairs based on phrase speech delay (>36 months) (Buxbaum *et al.*, 2001; Shao *et al.*, 2002). Both studies found increased support for linkage on chromosome 2q with MLSs of 2.86 and 3.32. An alternative approach was taken by Bradford *et al.*, who incorporated information on parental language phenotypes (Bradford *et al.*, 2001). When parent's history of language difficulties was taken into consideration they found that a subset of sib pairs where both probands had severe language delay were responsible for linkage signals on chromosome 7q and 13q. QTL analysis has also been carried out by Alarcon *et al.* using measures from the ADI for ''age at first word,'' ''age at first phrase'' and ''repetitive stereotyped behavior.'' Evidence for a QTL using a measure of ''age at first word'' overlapped with a QTL for repetitive behavior found on chromosome 7q (Alarcon *et al.*, 2002). The addition of 139 multiplex families to their study provided evidence for QTLs for ''age at first word'' at 3q13 and 17q and evidence for a language QTL on 16q which overlaps linkage found for SLI. The evidence for a language QTL on 7q diminished in the enlarged sample but ranking families according to language covariates gave support for a language-related locus. (Alarcon *et al.*, 2005).

Although these studies provide intriguing evidence that a susceptibility locus specifically relating to language exists for autism, their results are somewhat disparate and implicate already established susceptibility loci. Whether grouping sib pairs by language defines a particularly homogeneous group is debatable: language may be simply acting as a marker for autism severity. Attempting to utilise parental language data and measures from the ADI for analysis is also problematic given the unknown heritability of these traits. A rigorous QTL approach, as used by SLIC in their SLI genome screen, helps guard against the problems of unknown trait heritability and diminishing sample sizes when subsets are analysed. Suitable robust measures are not yet available for autism but exploratory analyses based on language

endophenotypes ascertained from the ADI-R have been performed as described above.

Cytogenetic studies for autism

The identification of *FOXP2* was aided by the characterization of the translocation of patient CS. Similarly, chromosomal abnormalities could give clues to the location of genes in multifactorial disease such as SLI and autism. Of particular interest in autism is the 15q11-q13 chromosomal region, which is most frequently reported as containing chromosomal abnormalities in autistic individuals and overlaps the Prader-Willi Angelman critical region (PWACR). There also appears to be a cluster of cytogenetic abnormalities in individuals with autism on chromosome 7, as shown in Figure 31.4 (visit www.chr7.org/ for more information on chromosome 7). Vorstman *et al.* have reviewed cytogenetic regions of interest (CROI) associated with an autistic phenotype. They identify regions of overlap between CROIs and linkage and association studies on chromosome 7, 15q11-q13 and 5p14 as well as novel regions with multiple case reports of cytogenetic abnormalities such as 22q11 and 22q13 (Vorstman *et al.*, 2006). However, only ~3–6% of autistic individuals possess chromosomal abnormalities and the cytogenetic evidence must be viewed with caution. The genetic pathways disrupted by these comparatively large-scale genomic abnormalities could provide information about unknown susceptibility gene variants, but may also be unrelated.

Association studies for autism

Numerous association studies have been carried out for autism both focused on candidate genes and using genome screen data (Vorstman *et al.*, 2006). In general, these studies suffer from a number of limitations: they have only relatively modest sample sizes of between 100–200 cases or trios, the *p*-values reported have only moderate significance ($P < 0.05$ but > 0.001) and insufficient SNPs have been genotyped to completely capture linkage disequilibrium across the candidate gene. There is however some evidence for association on chromosome 7q for autism and SLI, which support the linkage results, but these associations do not cluster closely together and have only a moderate significance (IMGSAC, 2001*b*; O'Brien *et al.*, 2003). Similarly, investigations into other candidate loci such as the serotonin transporter (5-HTT) gene on chromosome 17 and the gamma-aminobutyric acid (GABA) receptor on chromosome 15, have not been consistently replicated, casting doubt on the meaning of these associations (Lamb *et al.*, 2002). A systematic survey of a candidate region by association mapping with dense single nucleotide polymorphisms (SNP) maps, rather than narrowly focusing on a single gene, is now becoming more feasible and may present a useful way to localize susceptibility genes (Botstein and Risch, 2003).

Candidate gene investigations for SLI and autism

Following the identification of *FOXP2*, a number of DNA samples from language impaired patients have been screened for mutations (Bruce and Margolis, 2002; Meaburn *et al.*, 2002; Newbury *et al.*, 2002; Wassink *et al.*, 2002; Gauthier *et al.*, 2003; O'Brien *et al.*, 2003; Gong *et al.*, 2004; MacDermot *et al.*, 2005). In general, findings have been negative except for a single study investigating *FOXP2* in probands affected by verbal dyspraxia. Three novel coding changes were found including one missense stop codon in exon 7 that is likely to have functional significance (MacDermot *et al.*, 2005). Screening samples of individuals with autism, language impairment and idiopathic neurodegenerative movement disorders has detected some polymorphisms in the polyglutamine stretch of the gene and silent mutations (Newbury and Monaco, 2002). It is worth noting that a study of Chinese Han trios, affected by autism found some evidence for association with *FOXP2* but this has not be replicated in other samples and may represent a population specific effect (Gong *et al.*, 2004). It is possible that variants

Phenotype	Karyotype	Region of cytogenetic abnormality/breakpoint
Autism	46, XX, t(5;7) (q14;q35)	
Autism	46, XX t(7;13) (q32;q12)	
Autism	46, XY, t(7;21) (q32;q22.3)	
Autism	46, XY, dup(7) (q31.32;q34)	
Autism and developmental delay	46, XY, t(5;7) (q15;q31.32)	
Patient CS (FOXP2 disrupted)	46, XY, t(5;7) (q22;q31.2)	
Language impaired/ behavioral problems	46, XY, t(2;7) (p23;q31.3)	
Autism	46, XY, inv(7) (p12.2;q31.3)	
Autism	46, XY, t(7;13) (q31.3;q21)	
Autism	46, XX, t(7;11) (q31.2;q25)	
Autism	46, XY, inv(7) q22;q31.2)	
Autism	46, XY, t(6;7) (p11.2;q22)	
Autism, deafness and hyperactivity	46,XY, t(7;12) (q21.4;q15)	
Autism/PDDr	46, XY, del(7)(q21.13)	
Autism and mental retardation	46, X?, t(7;20) (q11.2;pp11.2)	
Autism	46, XY, t(1;7;21) (p22;q;q)	

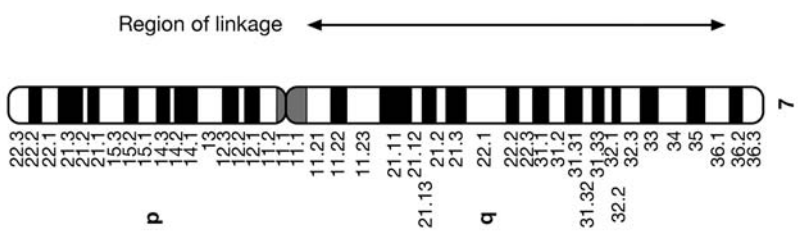

Figure 31.4 Cytogenetic cases associated with autism and region of linkage on chromosome 7. PDDr, pervasive developmental disorder.

exist in regulatory regions of *FOXP2* that have not yet been detected and affect speech and language, but it is unlikely that coding variants in the gene play a major role in the genetic susceptibility of autism or SLI.

A large number of genes have been screened for mutations in autistic individuals and a handful of missense variants detected. Few of these studies are powerful enough either to convincingly exclude a candidate gene or demonstrate a substantial

contribution to the genetic susceptibility to autism. Table 31.2 shows a selection of candidate gene studies for autism and information about replication studies. One of the few studies to detect numerous missense mutations in a single gene was for reelin (*RELN*) whose protein product is important for neuronal migration during brain development (Bonora *et al.*, 2003). Mice naturally null for the reelin protein have a characteristic reeling gait and possess neuroanatomical abnormalities in neuron positioning in the cerebral cortex, cerebellum and hippocampus. *RELN* maps beneath the peak of linkage reported by IMGSAC (International Molecular Genetic Study of Autism Consortium) at 7q22 and preferential transmission of a triplet repeat 5′ of the gene has been reported (Persico *et al.*, 2001). Thirteen missense mutations were identified in total by Bonora *et al.*, five of which were not present in controls. It remains difficult to assess the significance of these results since no behavioral phenotype could identify individuals with mutations. Furthermore, the numbers of families with missense mutations are insufficient to explain the strength of linkage detected on chromosome 7. Additional studies of reelin in independent autism samples have been carried out but the evidence they present is contradictory, as shown in Table 31.2.

Discussion

The identification of *FOXP2* is a significant step forward towards understanding the neural mechanisms underlying speech and language, but does not appear to have a wider role in the genetic aetiology of SLI or autism. It is likely that the R553H mutation is unique to the KE family and any other missense mutations represent single mutational events. Techniques available from Mendelian disease genetics were used to identify *FOXP2*, but these are less applicable to multigenic disorders. The potential complexity of many genes interacting in SLI and autism requires the collection of large well characterised samples and the implementation of sophisticated statistical analysis. Because the genetic mode of transmission for SLI and autism is unknown, and may vary between families, model based analysis is both less powerful than for Mendelian diseases and repeated analyses with different models can also lead to multiple testing issues. A non-parametric qualitative approach is more appropriate but necessitates defining individuals as affected or unaffected, which is often difficult when a broad clinical range of phenotypes exist. When suitable measures are available, a quantitative analysis is preferable because it is less restrictive in terms of diagnosis and is unbiased by model parameters.

Genome screens for SLI and autism have probably identified the key loci possessing genes with major or moderate effects but these genomic regions are broad and contain hundreds of genes. Association studies may more narrowly define these critical regions but require large sample sizes, affordable technology for SNP genotyping and robust statistical analysis as well as independent replication. Progress has been made toward affordable technology for SNP genotyping, but experiments can still be prohibitively expensive and strong debate continues among the statistics community about the most suitable analysis. Candidate gene studies for autism are numerous but few findings have been consistently replicated in independent samples. Furthermore, interpreting the significance of any mutations detected in candidate genes for multigenic disease is controversial and dependent on the underlying allelic spectrum for the disease at the locus (Reich and Lander, 2001; Pritchard and Cox, 2002).

Animal models for cognitive impairment have been developed using naturally occurring mutants, gene knockouts of candidate genes, surgical lesions and controlled chemical exposure. The success of these studies has been limited because of difficulties in interpreting phenotype characteristics. The serotonin transporter (5-HTT) knockout mouse has elevated anxiety and animal models for autism also exist (based on thalidomide exposure, Borna virus infection and oxytocin and vasopressin administration) but these are all likely to represent an

Table 31.2. Candidate gene studies for autism

Gene	Description of study and sample size	Results	Authors	Replication studies
WNT2	Coding sequence screened ($n = 135$) 2 SNPs typed ($n = 120$)	2 missense mutations Association with one SNP	(Wassink *et al.*, 2001)	Findings not replicated (McCoy *et al.*, 2002; Li *et al.*, 2004)
GRIK2	Coding sequence screened ($n = 33$) 3 SNPs typed ($n = 107$)	1 missense mutation found more often in autistic individuals	(Jamain *et al.*, 2002)	Postive association found (Shuang *et al.*, 2004)
RELN	5′ triplet repeat typed ($n = 165$)	Longer triplet repeats found more often in autistic individuals	(Persico *et al.*, 2001)	Association not replicated (Krebs *et al.*, 2002; Devlin *et al.*, 2004) Family based evidence for association found, possibly related to phrased speech (Zhang *et al.*, 2002) Association replicated (Skaar *et al.*, 2005)
	Coding sequence screened ($n = 55$) 2 SNPs and triplet repeat typed ($n = 169$)	13 missense mutations No evidence for association	(Bonora *et al.*, 2003)	Not done
cAMP-GEFII	Coding sequence screened ($n = 48$) 3 SNPs and insertion typed ($n = 164$)	4 missense mutations No evidence for association	(Bacchelli *et al.*, 2003)	Not done
HOXA1 & *HOXB1*	Coding sequence screened ($n = 57$)	3 missense mutations and one insertion	(Ingram *et al.*, 2000)	Findings not replicated (Devlin *et al.*, 2002; Li *et al.*, 2002; Talebizadeh *et al.*, 2002; Collins *et al.*, 2003; Romano *et al.*, 2003;

Table 31.2 (*cont.*)

Gene	Description of study and sample size	Results	Authors	Replication studies
				Gallagher *et al.*, 2004)
	1 SNP and insertion typed ($n = 57$)	Evidence for association		Positive association (Conciatori *et al.*, 2004)
NLGN3 & *NLGN4*	Coding sequence screened ($n = 148$)	2 missense mutations	(Jamain *et al.*, 2003)	Findings not replicated (Vincent *et al.*, 2004; Gauthier *et al.*, 2005; Ylisaukko-oja *et al.*, 2005; Blasi *et al.*, 2006)

oversimplified model for autism (Andres, 2002). The production of good animal models for specific cognitive impairments is obviously difficult and any advances will rely on the identification of susceptibility genes and neurological systems disrupted during development, as well as standardization of tools to analyse resulting phenotypes. Animal models for *FOXP2* will provide an insight into how much single gene models for complex cognitive phenotypes can aid our understanding of their molecular basis and whether this approach will be useful for multigenic disorders.

Research into speech and language disorders will benefit from developments in diagnostic measures that enable quantitative analysis of genotype–phenotype relationships. Additional pathophysiological and clinical studies may also help in localising structural and functional differences between controls and individuals with language impairment. Good hypothesis-driven research is necessary that accommodates advances in our knowledge of the variation present in the human genome and continued development of statistical analysis tools. Clear research goals are required for both molecular and clinical research to successfully locate the susceptibility genes causing the complex phenotype of language impairment.

Acknowledgements

We would like to thank Dianne Newbury, Simon Fisher and Janine Lamb for a critical reading of this chapter and assistance in creating figures. AJM is a Wellcome Principal Research Fellow.

Glossary

Clinical terms

Aspergers' syndrome – Autism with no language delay or general intellectual impairment and a range of milder or atypical manifestations.

Autism – The core triad of behaviors characteristic for autism includes: impairments in verbal and non-verbal communication, deficiencies in reciprocal social interaction and restricted activities and repetitive movements. Autism can be described as belonging to a spectrum of developmental disorders with overlapping clinical features.

Autistic regression – Apparently normal development followed by loss of communication and social skills.
Expressive language disorder – A disorder where a child has a lower than average proficiency in vocabulary, production of complex sentences and recall of words.
Macrocephaly – Head circumference greater than 1.88 standard deviations above normative data for age and sex.
Pervasive developmental disorder – Individuals may be diagnosed as having pervasive developmental disorder when the criteria for autism are not reached in all domains or where atypical symptoms are present.
Pragmatic language impairment – Inappropriate use of language within a social context.
Receptive language disorder – Impairment in the understanding of language.
Specific language impairment – Developmental language disorder that persists despite the absence of hearing loss, neurodevelopmental disorders and educational opportunity.

Technical terms

De novo – Occurring for the first time uniquely in one individual, i.e. not inherited.
FISH – Fluorescent in situ hybridization, a nucleic acid probe is labelled and hybridized to immobilized chromosome spreads where it binds to homologous sequences.
In silico analysis – Computer-based characterization of a gene using publicly available databases.
Non-parametric linkage analysis – Linkage analysis that relies on some specification of the components of a genetic model but not usually penetrance or degree of dominance. Components such as allele frequencies and recurrence risk may be specified.
Parametric linkage analysis – Linkage analysis that requires the specification of an underlying model including penetrance and degree of dominance.
Penetrance – The percentage of individuals of a specific genotype that express an expected phenotype, i.e. the likelihood that a person carrying a mutation will develop the characteristics caused by that mutation.
Quantitative trait – A variable non-dichotomous characteristic that depends on the effects of multiple genes and environmental influences.

REFERENCES

Alarcon, M., Cantor, R.M., Liu, J., Gilliam, T.C. and Geschwind, D.H. (2002). Evidence for a language quantitative trait locus on chromosome 7q in multiplex autism families. *Am J Hum Genet*, **70**, 60–71.

Alarcon, M., Yonan, A.L., Gilliam, T.C., Cantor, R.M. and Geschwind, D.H. (2005). Quantitative genome scan and ordered-subsets analysis of autism endophenotypes support language QTLs. *Mol Psychiatry*, **10**, 747–57.

Andres, C. (2002). Molecular genetics and animal models in autistic disorder. *Brain Res Bull*, **57**, 109–19.

Auranen, M., Vanhala, R., Varilo, T. *et al.* (2002). A genomewide screen for autism-spectrum disorders: evidence for a major susceptibility locus on chromosome 3q2527. *Am J Hum Genet*, **71**, 777–90.

Aylward, E.H., Minshew, N.J., Field, K., Sparks, B.F. and Singh, N. (2002). Effects of age on brain volume and head circumference in autism. *Neurology*, **59**, 175–83.

Bacchelli, E., Blasi, F., Biondolillo, M. *et al.* (2003). Screening of nine candidate genes for autism on chromosome 2q reveals rare non-synonymous variants in the cAMP-GEFII gene. *Mol Psychiatry*, In Press.

Badner, J.A. and Gershon, E.S. (2002). Regional meta-analysis of published data supports linkage of autism with markers on chromosome 7. *Mol Psychiatry*, **7**, 56–66.

Bailey, A., Le Couteur, A., Gottesman, I. *et al.* (1995). Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol Med*, **25**, 63–77.

Barrett, S., Beck, J.C., Bernier, R. *et al.* (1999). An autosomal genomic screen for autism. Collaborative linkage study of autism. *Am J Med Genet*, **88**, 609–15.

Bartlett, C.W., Flax, J.F., Logue, M.W. *et al.* (2002). A major susceptibility locus for specific language impairment is located on 13q21. *Am J Hum Genet*, **71**, 45–55.

Bartlett, C. W., Flax, J. F., Logue, M. W. *et al.* (2004). Examination of potential overlap in autism and language loci on chromosomes 2, 7, and 13 in two independent samples ascertained for specific language impairment. *Hum Hered*, **57**, 10–20.

Bartlett, C. W., Goedken, R. and Vieland, V. J. (2005). Effects of updating linkage evidence across subsets of data: reanalysis of the autism genetic resource exchange data set. *Am J Hum Genet*, **76**, 688–95.

Benasich, A. and Spitz, R. (1998). Insights from infants: temporal processing abilities and genetics contribute to language development. In Willems, G. and Whitmore, K. (eds.), *A neurodevelopmental approach to specific learning disorders*, pp. 191–210. London: MacKeith Press.

Bishop, D. V., Bishop, S. J., Bright, P. *et al.* (1999). Different origin of auditory and phonological processing problems in children with language impairment: evidence from a twin study. *J Speech Lang Hear Res*, **42**, 155–68.

Bishop, D. V., Laws, G., Adams, C. and Norbury, C. F. (2006). High heritability of speech and language impairments in 6-year-old twins demonstrated using parent and teacher report. *Behav Genet*.

Bishop, D. V. and Norbury, C. F. (2002). Exploring the borderlands of autistic disorder and specific language impairment: a study using standardised diagnostic instruments. *J Child Psychol Psychiatry*, **43**, 917–29.

Bishop, D. V., North, T. and Donlan, C. (1995). Genetic basis of specific language impairment: evidence from a twin study. *Dev Med Child Neurol*, **37**, 56–71.

Blasi, F., Bacchelli, E., Pesaresi, G. *et al.* (2006). Absence of coding mutations in the X-linked genes neuroligin 3 and neuroligin 4 in individuals with autism from the IMGSAC collection. *Am J Med Genet B Neuropsychiatr Genet*.

Bonora, E., Beyer, K. S., Lamb, J. A. *et al.* (2003). Analysis of reelin as a candidate gene for autism. *Mol Psychiatry*, In Press.

Botstein, D. and Risch, N. (2003). Discovering genotypes underlying human phenotypes: past successes for Mendelian disease, future approaches for complex disease. *Nat Genet*, **33** Suppl, 228–37.

Bradford, Y., Haines, J., Hutcheson, H. *et al.* (2001). Incorporating language phenotypes strengthens evidence of linkage to autism. *Am J Med Genet*, **105**, 539–47.

Bruce, H. A. and Margolis, R. L. (2002). *FOXP2*: novel exons, splice variants, and CAG repeat length stability. *Hum Genet*, **111**, 136–44.

Buxbaum, J. D., Silverman, J. M., Smith, C. J. *et al.* (2001). Evidence for a susceptibility gene for autism on chromosome 2 and for genetic heterogeneity. *Am J Hum Genet*, **68**, 1514–20.

Cantor, R. M., Kono, N., Duvall, J. A. *et al.* (2005). Replication of autism linkage: fine-mapping peak at 17q21. *Am J Hum Genet*, **76**, 1050–6.

Chakrabarti, S. and Fombonne, E. (2001). Pervasive developmental disorders in preschool children. *Jama*, **285**, 3093–9.

Collins, J. S., Schroer, R. J., Bird, J. and Michaelis, R. C. (2003). The HOXA1 A218G polymorphism and autism: lack of association in white and black patients from the South Carolina Autism Project. *J Autism Dev Disord*, **33**, 343–8.

Conciatori, M., Stodgell, C. J., Hyman, S. L. *et al.* (2004). Association between the *HOXA1* A218G polymorphism and increased head circumference in patients with autism. *Biol Psychiatry*, **55**, 413–19.

Cummings, C. J. and Zoghbi, H. Y. (2000). Fourteen and counting: unraveling trinucleotide repeat diseases. *Hum Mol Genet*, **9**, 909–16.

Dale, P. S., Price, T. S., Bishop, D. V. and Plomin, R. (2003). Outcomes of early language delay: I. Predicting persistent and transient language difficulties at 3 and 4 years. *J speech Lang Hear Res*, **46**, 544–60.

Dale, P. S., Simonoff, E., Bishop, D. V. *et al.* (1998). Genetic influence on language delay in two-year-old children. *Nat Neurosci*, **1**, 324–8.

Devlin, B., Bennett, P., Cook, E. H., Jr. *et al.* (2002). No evidence for linkage of liability to autism to HOXA1 in a sample from the CPEA network. *Am J Med Genet*, **114**, 667–72.

Devlin, B., Bennett, P., Dawson, G. *et al.* (2004). Alleles of a reelin CGG repeat do not convey liability to autism in a sample from the CPEA network. *Am J Med Genet B Neuropsychiatr Genet*, **126**, 46–50.

Enard, W., Przeworski, M., Fisher, S. E. *et al.* (2002). Molecular evolution of *FOXP2*, a gene involved in speech and language. *Nature*, **418**, 869–72.

Ferland, R. J., Cherry, T. J., Preware, P. O., Morrisey, E. E. and Walsh, C. A. (2003). Characterization of *Foxp2* and *Foxp1* mRNA and protein in the developing and mature brain. *J Comp Neurol*, **460**, 266–79.

Fisher, S. E., Vargha-Khadem, F., Watkins, K. E., Monaco, A. P. and Pembrey, M. E. (1998). Localisation of a gene implicated in a severe speech and language disorder. *Nat Genet*, **18**, 168–70.

Folstein, S. E., Santangelo, S. L., Gilman, S. E. *et al.* (1999). Predictors of cognitive test patterns in autism families. *J Child Psychol Psychiatry*, **40**, 1117–28.

Fombonne, E. (2005). Epidemiology of autistic disorder and other pervasive developmental disorders. *J Clin Psychiatry*, **66** Suppl 10, 3–8.

Fombonne, E. and Cook, E. H. (2003). MMR and autistic enterocolitis: consistent epidemiological failure to find an association. *Mol Psychiatry*, **8**, 133–4.

Gallagher, L., Hawi, Z., Kearney, G., Fitzgerald, M. and Gill, M. (2004). No association between allelic variants of HOXA1/HOXB1 and autism. *Am J Med Genet B Neuropsychiatr Genet*, **124**, 64–7.

Gauthier, J., Bonnel, A., St-Onge, J. *et al.* (2005). NLGN3/NLGN4 gene mutations are not responsible for autism in the Quebec population. *Am J Med Genet B Neuropsychiatr Genet*, **132**, 74–5.

Gauthier, J., Joober, R., Mottron, L. *et al.* (2003). Mutation screening of FOXP2 in individuals diagnosed with autistic disorder. *Am J Med Genet*, **118A**, 172–5.

Gong, X., Jia, M., Ruan, Y. *et al.* (2004). Association between the FOXP2 gene and autistic disorder in Chinese population. *Am J Med Genet B Neuropsychiatr Genet*, **127**, 113–16.

Hazlett, H. C., Poe, M., Gerig, G. *et al.* (2005). Magnetic resonance imaging and head circumference study of brain size in autism: birth through age 2 years. *Arch Gen Psychiatry*, **62**, 1366–76.

Hurst, J. A., Baraitser, M., Auger, E., Graham, F. and Norell, S. (1990). An extended family with a dominantly inherited speech disorder. *Dev Med Child Neurol*, **32**, 352–5.

IMGSAC (2001*a*). A genomewide screen for autism: strong evidence for linkage to chromosomes 2q, 7q, and 16p. *Am J Hum Genet*, **69**, 570–81.

IMGSAC (2001*b*). Further characterization of the autism susceptibility locus *AUTS1* on chromosome 7q. *Hum Mol Genet*, **10**, 973–82.

Ingram, J. L., Stodgell, C. J., Hyman, S. L. *et al.* (2000). Discovery of allelic variants of *HOXA1* and *HOXB1*: genetic susceptibility to autism spectrum disorders. *Teratology*, **62**, 393–405.

Jamain, S., Betancur, C., Quach, H. *et al.* (2002). Linkage and association of the glutamate receptor 6 gene with autism. *Mol Psychiatry*, **7**, 302–10.

Jamain, S., Quach, H., Betancur, C. *et al.* (2003). Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. *Nat Genet*, **34**, 27–9.

Krebs, M. O., Betancur, C., Leroy, S. *et al.* (2002). Absence of association between a polymorphic GGC repeat in the 5′ untranslated region of the reelin gene and autism. *Mol Psychiatry*, **7**, 801–4.

Lai, C. S., Fisher, S. E., Hurst, J. A. *et al.* (2000). The SPCH1 region on human 7q31: genomic characterization of the critical interval and localization of translocations associated with speech and language disorder. *Am J Hum Genet*, **67**, 357–68.

Lai, C. S., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F. and Monaco, A. P. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*, **413**, 519–23.

Lai, C. S., Gerrelli, D., Monaco, A. P., Fisher, S. E. and Copp, A. J. (2003). *FOXP2* expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. *Brain*.

Lai, C. S., Gerrelli, D., Monaco, A. P., Fisher, S. E. and Copp, A. J. (2003). *FOXP2* expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. *Brain*, **126**, 2455–62.

Lainhart, J. E., Ozonoff, S., Coon, H. *et al.* (2002). Autism, regression, and the broader autism phenotype. *Am J Med Genet*, **113**, 231–7.

Lamb, J. A., Barnby, G., Bonora, E. *et al.* (2005). Analysis of IMGSAC autism susceptibility loci: evidence for sex limited and parent of origin specific effects. *J Med Genet*, **42**, 132–7.

Lamb, J. A., Parr, J. R., Bailey, A. J. and Monaco, A. P. (2002). Autism: in search of susceptibility genes. *Neuromolecular Med*, **2**, 11–28.

Law, J., Boyle, J., Harris, F., Harkness, A., Nye, C. (1998). Screening for speech and language delay: a systematic review of the literature. *Health Technol Assess*, **2**, 1–184.

Li, J., Nguyen, L., Gleason, C. *et al.* (2004). Lack of evidence for an association between WNT2 and RELN polymorphisms and autism. *Am J Med Genet B Neuropsychiatr Genet*, **126**, 51–7.

Li, J., Tabor, H. K., Nguyen, L. *et al.* (2002). Lack of association between HoxA1 and HoxB1 gene variants and autism in 110 multiplex families. *Am J Med Genet*, **114**, 24–30.

Lord, C. and Bailey, A. (2002). Autism Spectrum Disorders. In Rutter, M. and Taylor, E. (eds.), *Child and adolescent psychiatry*, p. 1209. *Vol 1*. Oxford: Blackwell Publishers.

MacDermot, K. D., Bonora, E., Sykes, N. *et al.* (2005). Identification of *FOXP2* truncation as a novel cause of

developmental speech and language deficits. *Am J Hum Genet*, **76**, 1074–80.

Madsen, K. M., Hviid, A., Vestergaard, M. *et al.* (2002). A population-based study of measles, mumps, and rubella vaccination and autism. *N Engl J Med*, **347**, 1477–82.

Marcus, G. F. and Fisher, S. E. (2003). *FOXP2* in focus: what can genes tell us about speech and language? *Trends Cogn Sci*, **7**, 257–62.

McCoy, P. A., Shao, Y., Wolpert, C. M. *et al.* (2002). No association between the WNT2 gene and autistic disorder. *Am J Med Genet*, **114**, 106–9.

Meaburn, E., Dale, P. S., Craig, I. W. and Plomin, R. (2002). Language-impaired children: No sign of the *FOXP2* mutation. *Neuroreport*, **13**, 1075–7.

Newbury, D. F., Bonora, E., Lamb, J. A. *et al.* (2002). FOXP2 is not a major susceptibility gene for autism or specific language impairment. *Am J Hum Genet*, **70**, 1318–27.

Newbury, D. F. and Monaco, A. P. (2002). Molecular genetics of speech and language disorders. *Curr Opin Pediatr*, **14**, 696–701.

Nishimura, D. Y., Searby, C. C., Alward, W. L. *et al.* (2001). A spectrum of *FOXC1* mutations suggests gene dosage as a mechanism for developmental defects of the anterior chamber of the eye. *Am J Hum Genet*, **68**, 364–72.

O'Brien, E. K., Zhang, X., Nishimura, C., Tomblin, J. B. and Murray, J. C. (2003). Association of specific language impairment (SLI) to the region of 7q31. *Am J Hum Genet*, **72**, 1536–43.

Persico, A. M., D'Agruma, L., Maiorano, N. *et al.* (2001). Reelin gene alleles and haplotypes as a factor predisposing to autistic disorder. *Mol Psychiatry*, **6**, 150–9.

Pickles, A., Bolton, P., Macdonald, H. *et al.* (1995). Latent-class analysis of recurrence risks for complex phenotypes with selection and measurement error: a twin and family history study of autism. *Am J Hum Genet*, **57**, 717–26.

Pickles, A., Starr, E., Kazak, S. *et al.* (2000). Variable expression of the autism broader phenotype: findings from extended pedigrees. *J Child Psychol Psychiatry*, **41**, 491–502.

Pritchard, J. K. and Cox, N. J. (2002). The allelic architecture of human disease genes: common disease-common variant.or not? *Hum Mol Genet*, **11**, 2417–23.

Rapin, I. (1998). Understanding childhood language disorders. *Curr Opin Pediatr*, **10**, 561–6.

Reich, D. E. and Lander, E. S. (2001). On the allelic spectrum of human disease. *Trends Genet*, **17**, 502–10.

Romano, V., Cali, F., Mirisola, M. *et al.* (2003). Lack of association of *HOXA1* and *HOXB1* mutations and autism in Sicilian (Italian) patients. *Mol Psychiatry*, **8**, 716–17.

Rutter, M. (2005). Incidence of autism spectrum disorders: changes over time and their meaning. *Acta Paediatr*, **94**, 2–15.

Shao, Y., Raiford, K. L., Wolpert, C. M. *et al.* (2002). Phenotypic homogeneity provides increased support for linkage on chromosome 2 in autistic disorder. *Am J Hum Genet*, **70**, 1058–61.

Shuang, M., Liu, J., Jia, M. X. *et al.* (2004). Family-based association study between autism and glutamate receptor 6 gene in Chinese Han trios. *Am J Med Genet B Neuropsychiatr Genet*, **131**, 48–50.

Skaar, D. A., Shao, Y., Haines, J. L. *et al.* (2005). Analysis of the RELN gene as a genetic risk factor for autism. *Mol Psychiatry*, **10**, 563–71.

SLIC (2002). A genomewide scan identifies two novel loci involved in specific language impairment. *Am J Hum Genet*, **70**, 384–98.

SLIC (2004). Highly significant linkage to the SLI1 locus in an expanded sample of individuals affected by specific language impairment. *Am J Hum Genet*, **74**, 1225–38.

Tager-Flusberg, H. (2000). The challenge of studying language development in children with autism. In LM, NBR (eds.), *Methods for studying language production*, pp. 313–31. Mahwah, NJ: Erlbaum.

Takahashi, K., Liu, F. C., Hirokawa, K. and Takahashi, H. (2003). Expression of *Foxp2*, a gene involved in speech and language, in the developing and adult striatum. *J Neurosci Res*, **73**, 61–72.

Talebizadeh, Z., Bittel, D. C., Miles, J. H. *et al.* (2002). No association between *HOXA1* and *HOXB1* genes and autism spectrum disorders (ASD). *J Med Genet*, **39**, e70.

Tomblin, J. B. and Buckwalter, P. R. (1998). Heritability of poor language achievement among twins. *J Speech Lang Hear Res*, **41**, 188–99.

Tomblin, J. B., Hafeman, L. L. and O'Brien, M. (2003). Autism and autism risk in siblings of children with specific language impairment. *Int J Lang Commun Disord*, **38**, 235–50.

Trikalinos, T. A., Karvouni, A., Zintzaras, E. *et al.* (2006). A heterogeneity-based genome search meta-analysis for autism-spectrum disorders. *Mol Psychiatry*, **11**, 29–36.

Vargha-Khadem, F., Watkins, K., Alcock, K., Fletcher, P. and Passingham, R. (1995). Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. *Proc Natl Acad Sci USA*, **92**, 930–3.

Vargha-Khadem, F., Watkins, K. E., Price, C. J. *et al.* (1998). Neural basis of an inherited speech and language disorder. *Proc Natl Acad Sci USA*, **95**, 12695–700.

Vincent, J. B., Kolozsvari, D., Roberts, W. S. *et al.* (2004). Mutation screening of X-chromosomal neuroligin genes: no mutations in 196 autism probands. *Am J Med Genet B Neuropsychiatr Genet*, **129**, 82–4.

Vorstman, J. A., Staal, W. G., van Daalen, E. *et al.* (2006). Identification of novel autism candidate regions through analysis of reported cytogenetic abnormalities associated with autism. *Mol Psychiatry*, **11**, 1, 18–28.

Wassink, T. H., Piven, J., Vieland, V. J. *et al.* (2001). Evidence supporting WNT2 as an autism susceptibility gene. *Am J Med Genet*, **105**, 406–13.

Wassink, T. H., Piven, J., Vieland, V. J. *et al.* (2002). Evaluation of *FOXP2* as an autism susceptibility gene. *Am J Med Genet*, **114**, 566–9.

Ylisaukko-oja, T., Rehnstrom, K., Auranen, M. *et al.* (2005). Analysis of four neuroligin genes as candidates for autism. *Eur J Hum Genet*, **13**, 1285–92.

Zhang, H., Liu, X., Zhang, C. *et al.* (2002). Reelin gene alleles and susceptibility to autism spectrum disorders. *Mol Psychiatry*, **7**, 1012–17.

32

Common forms of visual handicap

Alan Wright

Introduction

The evolution, comparative structure and function of the eye have attracted the attention of many scientists, including Isaac Newton, who first showed that light could be split into different wavelengths, and later Charles Darwin, for whom the eye presented an interesting test of the evolutionary paradigm. Newton laid the foundations for the trichromatic theory of vision (Young, 1802), whereby different wavelengths of light are perceived by three distinct receptors within the retina, with overlapping sensitivities, now known to be the short-, medium- and long-wavelength cone opsins. The eye interacts with the environment in the most direct way, since it is constantly bombarded with electromagnetic radiation of many wavelengths and yet it specifically responds only to those in the 400–700 nm range. These wavelengths correspond closely to the solar spectrum measured below the surface of the sea, where the earliest visual systems are thought to have evolved (McIlwaine, 1996). An otherwise rare member of the carotenoid family, 11-cis retinal, an isomer of vitamin A aldehyde, has the important property of changing shape on absorbing a photon of light. This molecule combines with the different opsin apoproteins, which are members of the G-protein coupled receptor superfamily, to form the visual photopigment in all multicellular animals. In this way, light serves to generate a neural signal but at the same time renders the eye vulnerable to oxidative damage, which is a major factor in at least two of the major causes of global blindness.

Age-related cataract

The commonest cause of visual impairment globally is cataract, which is defined as opacity or loss of transparency within the crystalline lens, leading to reduced vision (Figure 32.1). This occurs when the refractive index of the lens varies over distances similar to the wavelength of light (Delaye and Tardieu, 1983). The causes of cataract are many and varied but include changes in lens cell architecture or in its protein constituents. For example, light scattering can result from aggregates in the size range of 100 nm or greater.

About 37 million people worldwide are blind and over 160 million visually impaired (Resnikoff *et al.*, 2004). The proportion of the population with severe visual impairment rises steeply with age, but the causes differ between continents and countries, varying from 0.1–0.2% in industrialized countries to over 1% in sub-Saharan Africa (Resnikoff *et al.*, 2004). About 90% of the world's blind live in developing countries, where the major cause of blindness is cataract. Globally, cataract accounts for 48% of all blindness, almost four times more than any other disorder. There is no known means of cataract prevention but in industrialized countries, cataract removal is a relatively simple and successful operation. The prevalence in a middle class white United Kingdom (UK) population was 49% of

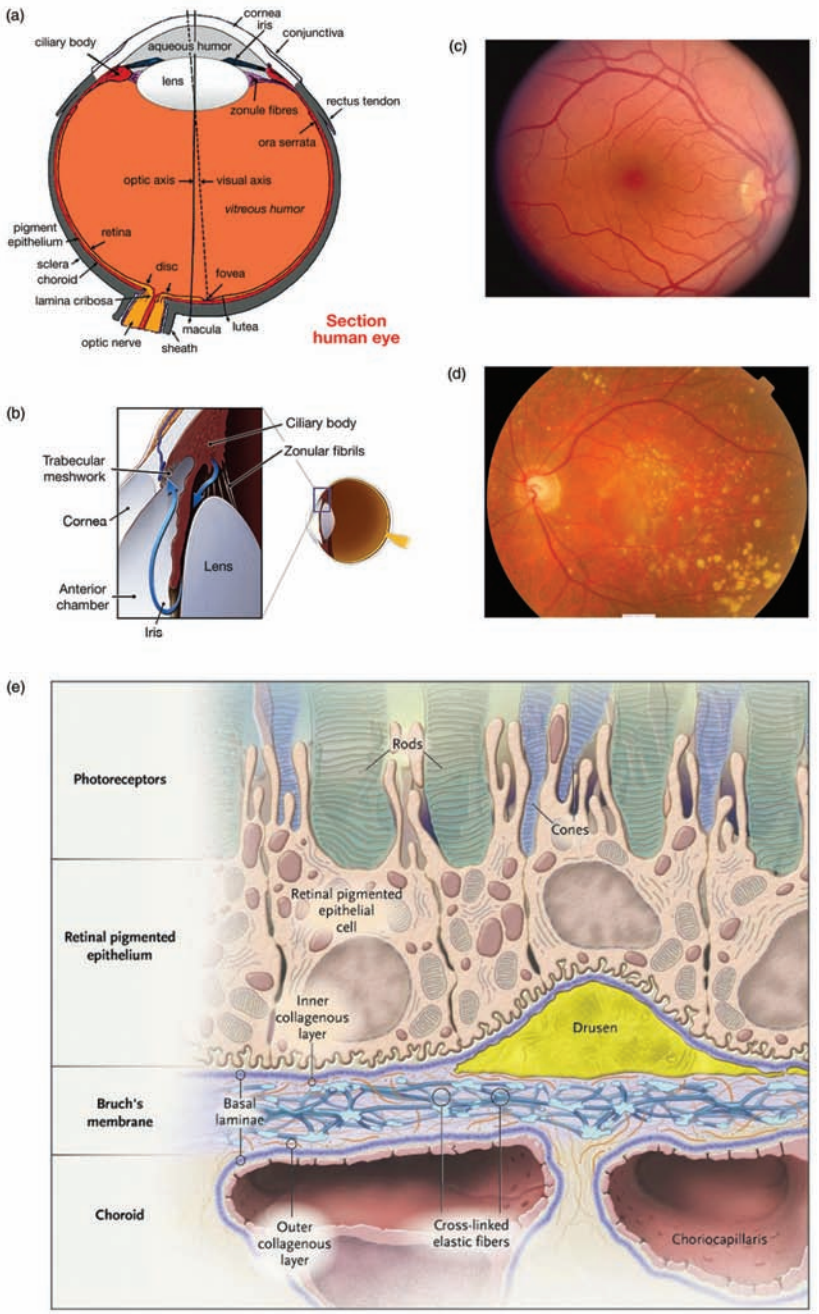

Figure 32.1 Anatomical features of the human eye. (a) Human eye showing the main anatomical features referred to in the text. The visual axis is the line connecting the object of fixation and the fovea, so that light is neither reflected nor refracted. This differs from the optic axis, which passes through the centre of the cornea and lens, which are rarely co-axial. The limits of the macula lutea are shown (arrows) flanking the fovea, the region of highest visual acuity

those in the age range 75–84 years, rising to 71% of those over age 85 years, with slightly higher rates in females than males (Reidy *et al.*, 1998). In the UK, it is the commonest surgical operation, with over 160 000 extractions per year compared with a global figure of about 10 million per year. As a result of treatment, cataract-related blindness is declining in many countries, despite their increasingly elderly populations.

There are three main types of cataract: nuclear, cortical and posterior subcapsular. These do not readily correspond to distinct etiological groups, although nuclear cataracts are the most common form of age-related cataract, followed by cortical cataracts, which are most common in those under age 75 years. Besides age, one of the major risk factors for cataract is exposure to ultraviolet B light, which has been proposed to explain about 10% of cortical cataracts in some populations. In one study, a doubling of exposure to UV light increased the risk of cortical cataract by 60% (Taylor *et al.*, 1998). The lens is vulnerable to damage from solar radiation, which generates reactive oxygen species and may lead to protein cross-linking and aggregation, lipid peroxidation and DNA damage, all of which can lead to cell death (Hejtmancik *et al.*, 2001). The lens has powerful antioxidants and repair capacity but these defences are insufficient to prevent age-associated change. Oxidation of lens proteins is closely associated with cataract formation (Hejtmancik *et al.*, 2001) but there is only weak or conflicting evidence for protection by antioxidant nutrients.

Other risk factors for cataract include cigarette smoking, which may account for as much 16–17% of nuclear cataract in some populations, diabetes mellitus, hypertension, obesity, corticosteroids, female gender and conditions such as uveitis, degenerative disorders of the retina, myopia, ocular trauma and surgery (Congdon and Taylor, 2003).

Recently, evidence has been presented for genetic influences in age-related cataract. Population-based studies show two- to threefold relative risks of cataract for those with an affected first-degree relative, although this does not distinguish between shared environmental and genetic factors. Twin studies suggest that the greater similarity of identical versus non-identical twins for both nuclear and cortical cataracts cannot be due to common environment but rather is due to both heritable and non-shared environmental differences (Hammond *et al.*, 2000; 2001). A large twin study of 506 female twin pairs from the UK population, aged 50–79 years, showed that the heritability of age-related nuclear cataract was 0.48, only slightly lower than for cortical cataract (0.53–0.58) (Hammond *et al.*, 2000). Age accounted for 38% of the variance in nuclear cataract, genetic variation for 48%, while unique

(strictly, the centre of the fovea (foveola) shows maximum acuity). (redrawn from remort.wz.cz/retina/II/images/DRAWEYE.png). (b) Drainage of aqueous humor, which is formed by the ciliary body, into the trabecular meshwork. Drainage abnormalities are implicated in glaucoma (redrawn from www.ahaf.org/glaucoma/about/aqueousflowBorder.jpg). (c) Normal human retinal fundus photograph showing the macula lutea, the small pigmented spot at the centre of the posterior retina, which is damaged in age-related macular degeneration. The term macula commonly refers to the area of central retina (area centralis) lying between the superior and inferior temporal retinal arteries (vascular arcades). The underlying retinal pigment epithelium and Bruch's membrane are not seen. (d) Human retinal fundus photograph from a patient with age-related macular degeneration (AMD) showing abnormal extracellular deposits or drusen, one of the clinical hallmarks of AMD, in the macular region. (e) Schematic diagram of sub-RPE deposits in age-related macular degeneration (AMD). Macular cone and rod photoreceptors are shown inter-digitating with the villous extensions of the underlying retinal pigment epithelium (RPE), Bruch's membrane and choroidal blood vessels. A single extracellular deposit or druse is shown between RPE and Bruch's membrane. See text for further details. Image from Johnson and Anderson (2004) *N Engl J Med* **351**, 320–2.

environmental factors explained only 14% of the variance (Hammond *et al.*, 2000). These estimates were based on the higher correlation (r) for the presence of nuclear cataract within identical (monozygotic, MZ) twins ($r = 0.90$) compared with non-identical (dizygotic, DZ) twins ($r = 0.57$). The high correlation within DZ twins suggests the importance of shared environmental factors, including age.

In age-related cortical cataract, genetic factors showing either additive (0–20% of variance) or dominance (38–53% of variance) effects contributed about half (53–58%) of the variance and unique (non-shared) environment explained 26–37% of the variance, the remainder being due to age (Hammond *et al.*, 2001). These studies illustrate the value of the ''perfect natural experiment'' in which MZ and DZ twin pairs are compared to elucidate the relative importance of genetic and environmental sources of variability (Martin *et al.*, 1997).

Racial differences in cataract prevalence have also been reported. The odds of African-Americans having cortical cataracts are four times greater than for Whites living in the same community (West *et al.*, 1998). However, both nuclear (odds ratio 2.1) and posterior subcapsular (odds ratio 2.5) cataracts are more common in Caucasians (West *et al.*, 1998). Indians who have moved to the UK and share similar exposure to solar radiation also have a higher prevalence of cataract than the native British population, although dietary and other environmental factors cannot be excluded.

A large number of genes have been implicated in congenital cataract, which accounts for about one-tenth of childhood blindness worldwide. There are over 150 mostly syndromic disorders associated with human congenital cataract listed in the Online Mendelian Inheritance in Man (OMIM) catalogue (www.ncbi.nlm.nih.gov/omim). A smaller number are associated with isolated (non-syndromic) congenital cataract. Specifically, about 30 congenital cataract genes have been mapped, 15 of which have been identified (Table 32.1). Many of these genes have corresponding mutations among the approximately 40 cataract genes in the laboratory mouse (Graw, 2004). The prevalence of congenital cataract is about 1 in 3000 and one-third to one-half have a genetic basis (Graw, 2004). A common genetic cause of congenital cataract is mutation in one or other of the genes producing stable, water-soluble lens crystallins, which account for almost 90% of lens proteins and are essential for lens transparency. Maintaining the solubility of crystallins is critical for lens transparency, and point mutations which only alter their solubility, rather than their structure, are capable of causing cataract (Hejtmancik and Kantorow, 2004). Other genetic causes include mutations in gap junction proteins, water channels, membrane transporters, cytoskeletal proteins and transcription factors (Table 32.1) (Hejtmancik and Kantorow, 2004; Reddy *et al.*, 2004). Mutations in most of these genes cause autosomal dominant forms of congenital cataract.

It has been suggested that more subtle mutations affecting some of the same congenital cataract genes could influence susceptibility to age-related cataract, although this remains to be demonstrated (Hejtmancik and Kantorow, 2004). It is known that crystallins are subject to a wide variety of post-translational modifications which increase with age, including proteolysis, oxidation of sulphydryl groups, deamidation of asparagine and glutamine residues, phosphorylation, non-enzymatic glycosylation and other changes (Hejtmancik *et al.*, 2001). Maintaining their solubility and transparency throughout life is therefore a challenge and subtle genetic changes could contribute to insolubility. α-Crystallins are known to bind and help maintain the solubility of age-modified βγ-crystallins, which reduces light scattering. However, α-crystallins do not renature their binding partners, so that the resultant complexes increase in size and deplete the available soluble crystallins so that light scattering does occur (Reddy *et al.*, 2004). A number of congenital cataract genes have been implicated in juvenile or progressive forms of cataract. Juvenile cataracts can result from

Table 32.1. Genes associated with monogenic forms of congenital cataract (see text)

Gene name	Chromosomal location	Protein function (mutations)
Crystallins		
CRYAA	21q22.3	Subunit of α-crystallin, which has small heat shock protein and chaperone activity (~35% of lens crystallins) (*R116C, R49C* (dominant) and *W9X* (recessive))
CRYAB	11q22–22.3	Subunit of α-crystallin, which has small heat shock protein and chaperone activity (~35% of lens crystallins) (Frameshift, R120G)
CRYBA1	17q11.1–12	β-crystallin subunit, the most abundant water-soluble protein in the lens (Splice site)
CRYBB1	22q11	Subunit of β-crystallin, the most abundant water-soluble protein in the lens (~55% of lens crystallins) (*G220X*)
CRYBB2	22q11.2	Subunit of β-crystallin, the most abundant water-soluble protein in the lens (~55% of lens crystallins) (Q155X)
CRYGC	2q33–35	Subunit of γ-crystallin; abundant water-soluble lens protein; similar in structure to β-crystallin (~10% of lens crystallins) (T5P, insertion)
CRYGD	2q33–35	Subunit of γ-crystallin; abundant water-soluble lens protein; similar in structure to β-crystallin (~10% of lens crystallins) (T5P, insertion)
Membrane transport proteins		
CX50	1q21–25	Component of low resistance gap junction channels between lens fibre cells; necessary for flow of ions and metabolites between cells (S88P, E48K, I247M)
CX46	13q11–13	Component of low resistance gap junction channels between lens fibre cells; necessary for flow of ions and metabolites between cells (N63S, Frameshift, P187L)
AQP0 (MIP)	12q12	Member of aquaporin family; mediates rapid flow of water across membranes; abundant membrane protein in lens fibre cells (T138R, E134G)
Cytoskeletal proteins		
BFSP2	3q21.2	A component of beaded filaments, a cytoskeletal structure restricted to lens fibre cells (R287W, delE233)
Transcription factors		
MAF	16q23.2	Basic leucine zipper transcription factor which binds MAF responsive elements in in the promoters of PITX3 and crystallin genes (*R288P*)
PITX3	10q24–25	Paired class of homeodomain transcription factor (S13N)
HSF4	16q22	Heat shock transcription factor; regulates small and large heat shock proteins (A20D, I87V, L115P, R120C)
Unknown function		
LIM2 (MP19)	19q13.4	Second most abundant protein of the lens fibre cell membrane; function unknown (F105V, recessive)

mutation in the *BFSP2* gene and progressive juvenile cataracts can occur with mutations in the *γC-crystallin* and *aquaporin 0* (MIP) genes (Conley *et al.*, 2000; Francis *et al.*, 2000; Ren *et al.*, 2000). Age-related cataract, defined as onset after the age of 40 years (Hejtmancik and Kantorow, 2004), can be caused by a number of rare monogenic disorders. *Galactokinase* (*GALKI*) is one of three genes implicated in galactosemia, all of which can cause congenital cataracts. Mild mutations in *GALK1* result in age-related cataract, as a result of galactitol accumulation, which causes osmotic swelling of lens fibre cells. For example, the Osaka variant (*Ala198Val*) of galactokinase,

which is found exclusively in Japanese and other eastern populations, is present in 8% of Japanese with age-associated cataract compared with 4% of the general population ($P<0.02$)(Okano *et al.*, 2001). Diabetes mellitus is a known risk factor for cataract, which here also results from polyol accumulation. Hyperglycemia results in greatly increased glucose uptake in tissues where uptake is independent of insulin, including lens, retina, kidney and peripheral nerves (Harding, 1991). As a result, the glucose flux through the polyol pathway is increased tenfold. The excess glucose is reduced by aldose reductase to sorbitol, which accumulates and becomes hydrated in lens fibres and epithelium, leading to osmotic stress, a major factor in diabetic cataract. Susceptibility to diabetic cataract has been associated with a specific polymorphism upstream of the aldose reductase gene (Lee *et al.*, 2001), although this remains to be replicated. Finally, common variation in the glutathione-S-transferase gene has been associated with age-related cataract but this has not been consistently replicated (Hejtmancik and Kantorow, 2004).

The aim of identifying individual genes associated with age-related cataract in families has recently been furthered by a genome-wide linkage scan (Iyengar *et al.*, 2004). A series of 224 sib pairs, each with quantitative scores reflecting the extent of cortical cataract was analysed for linkage to microsatellite markers. This identified two chromosomal regions (1p35, 6p12-q12) showing significant evidence for linkage. The most significant linked region, on chromosome 6, contained about 267 genes, posing a formidable problem for gene identification (Iyengar *et al.*, 2004). These findings remain to be replicated, so that the identification of age-related cataract genes remains at an early stage, with much still to be learnt.

Glaucoma

Glaucoma is a diverse group of progressive optic nerve disorders (neuropathies) which accounts for 12% of global blindness, second only to cataract (Resnikoff *et al.*, 2004). In Western countries, where cataract-related blindness is substantially less common (about 5% of blindness), glaucoma is also the second most common cause, accounting for 18% of all blindness (Resnikoff *et al.*, 2004). The diagnosis of adult-onset glaucoma is primarily based on characteristic structural abnormalities of the optic nerve head (optic disc; Figure 32.1) together with functional abnormalities in the associated field of vision (Johnstone and Quigley, 2003). In the majority of cases, raised intra-ocular pressure (IOP $>$99.5th percentile) is also present, but this is no longer regarded as necessary for the diagnosis, since one-sixth to one-third of patients have normal IOP. Raised IOP is certainly a major risk factor for glaucoma and genetic studies (see below) suggest that in some families it is a primary etiological factor.

Glaucoma is classified into several subtypes, of which primary open-angle glaucoma (POAG) and primary angle-closure glaucoma (PACG) are the most common, with secondary glaucoma, due to other conditions, and congenital glaucoma, being much less common. The prevalence of POAG and PACG varies markedly between ethnic groups. The prevalence of POAG in white Europeans over the age of 40 years is 2.5–3%, and it is four times as common in Africans, and about twice as common globally as PACG. However, the rate of blindness is higher in PACG, which is more common in China and South-East Asia than in Europe, and so it may cause more blindness than POAG (Quigley, 2004). The Japanese are an Asian exception, with more POAG than PACG, but the disease mechanism may differ from Europeans, due to population differences in IOP and myopia.

Primary open angle glaucoma

POAG is a slowly progressive disorder of unknown cause in which optic neuropathy develops in the context of abnormalities of the central field of vision, with or without raised IOP. The fluid bathing the anterior chamber of the eye

(Figure 32.1), which lies between the cornea and iris/lens, is called the aqueous humor. It is continuously secreted by epithelial cells of the ciliary body and drains into the ''filtration angle'' between the cornea and iris, which contains a porous structure called the trabecular meshwork. The trabecular meshwork communicates with the blood stream via the canal of Schlemm and increased resistance in outflow through this structure may contribute to raised IOP and optic neuropathy. The trabecular meshwork in POAG patients is not structurally abnormal, in contrast to the situation in PACG. Since raised IOP is often absent in POAG, other factors, including disorders of the blood vessels supplying the optic nerve head, or of the retinal ganglion cells (the source of the optic nerve fibres), have been proposed to be of greater importance (Quigley, 2004).

Risk factors in POAG include ones related to onset and others related to progression. Age is a major risk factor, with prevalence rising steeply from <0.1% before age 30 years to 5–10% or greater in the elderly (Johnstone and Quigley, 2003). A major risk factor for both disease onset and progression is IOP, which in Europeans has a mean of 16 mm Hg and standard deviation of 2.5 mm Hg, but since the distribution is non-normal, 5–7% (rather than 2.5%) of the population has IOP greater than 21 mm Hg, which is often regarded as the upper limit of normal. The risk of glaucoma rises with increasing IOP, following a shallow exponential curve with no evident threshold. In addition, there is a normal diurnal variation of 5 mm Hg that can reach even higher levels in glaucoma. The higher the IOP, the greater the risk of optic nerve damage, and lowering of IOP has been shown to be beneficial in glaucoma whether or not it is elevated (Heijl *et al.*, 2002). Other risk factors for POAG include myopia (short sight), corticosteroid use, which elevates IOP in susceptible individuals, and finally family history.

It has long been recognised that glaucoma clusters within families. One-third to one-half of glaucoma and ocular hypertensive (raised IOP) patients have a family history of the same disorder (Williams-Lyn *et al.*, 2000). First-degree relatives of POAG patients have a 7–10-fold increased risk compared with the general population (Tielsch *et al.*, 1991). The mode of inheritance is multifactorial, although rare families with an autosomal dominant mode of inheritance can be found, often with onset at a young age, either in juveniles or young adults. No large-scale twin or adoption studies have been reported in POAG, so that familial resemblance due to shared environmental factors is not excluded. Further evidence for a genetic influence in POAG however comes from the fourfold increased risk in those of black African origin, which applies equally to Tanzanian villagers and East Baltimore residents, despite radically different environments (Buhrmann *et al.*, 2000).

Seven genes (*GLC1A-G*) have been mapped in POAG families showing apparently monogenic forms of glaucoma, two of which have been identified (Stone *et al.*, 1997; Rezaie *et al.*, 2002; Gong *et al.*, 2004). The *myocilin* (*MYOC*) gene was the first to be identified by positional cloning, following linkage mapping of the *GLC1A* locus to chromosome 1q24.3-q25.2 in autosomal dominant juvenile-onset open angle glaucoma (JOAG) families (Stone *et al.*, 1997). Over 70 disease-causing *MYOC* mutations have since been identified in POAG patients with both juvenile and adult onset (Gong *et al.*, 2004). Disease-causing *MYOC* mutations have been found in unselected POAG patients at a similar frequency in different ethnic groups – 4% of POAG or normal-tension glaucoma patients of White origin, 3.3% of black African origin and 4.4% of Asian origin (Gong *et al.*, 2004). The majority of these mutations are not found in controls, but two are also present at low frequency in controls. The *Gln368Stop* mutation is found in 2–3% of Caucasian POAG patients and about 0.3–0.4% of controls, contributing almost one-half of all White POAG mutations, although it is very rare in other ethnic groups (Stone *et al.*, 1997; Gong *et al.*, 2004). There has been uncertainty as to whether this variant is disease-causing or disease-modifying, since it has been found in

unaffected controls. Family studies show that the relative risk of POAG in a *Gln368Stop* carrier is 13.1, consistent with this being a major disease-risk allele, although the age-of-onset tends to be later than with other POAG variants (Gong *et al.*, 2004). The other common mutation, *Arg46Stop*, is found in about 1% of Asian POAG patients, but not in other ethnic groups. In family studies, it confers a relative risk of 5.4 ($P = 0.1$) for developing POAG, suggesting that it does influence POAG risk, although the numbers are small and the confidence limits wide (Gong *et al.*, 2004).

The function of myocilin is uncertain. It was originally found in trabecular meshwork cells when induced by glucocorticoids (Nguyen *et al.*, 1998). It is widely expressed although is most abundant in the iris, ciliary body and trabecular meshwork (Aroca-Aguilar *et al.*, 2005). It is a 55 kDa protein with a signal peptide, which may or may not be active, and an amino (N)-terminal double stranded coiled-coil leucine zipper domain, which is a protein structure well suited to protein interactions and is involved in myocilin dimerisation. This is followed by a domain of unknown function and then a distinct carboxyl (C)-terminal domain similar to olfactomedin, an extracellular matrix protein of unknown function. Mature myocilin forms multimers and the normal protein is secreted into the trabecular extracellular matrix (ECM) where it appears to interact with various ECM components. The majority of POAG mutations occur in the olfactomedin domain, at least some of which are known to inhibit endoproteolytic cleavage of the protein into a secreted 35 kDa protein that is present in aqueous humour (Aroca-Aguilar *et al.*, 2005). The mutant protein is retained in the endoplasmic reticulum and in transfected cells forms insoluble aggregates. This may lead to loss of trabecular meshwork cells and raised IOP although the pathway between genotype and phenotype is complex.

A second reported POAG gene is optineurin (*OPTN*, *GLC1E*) (Rezaie *et al.*, 2002), however, follow-up studies indicate that *OPTN* mutation is a rare cause of both POAG and normal tension glaucoma, probably accounting for less than 0.1% of all POAG (Wiggs *et al.*, 2003; Alward *et al.*, 2003). Another gene, *CYP1B1*, appears to be a POAG modifier gene. Affected members of an autosomal dominant POAG family carrying a *Gly399Val* mutation in *MYOC*, together with an *Arg368His* mutation in the *CYP1B1* gene (previously associated with primary congenital glaucoma, PCG), had juvenile-onset open angle glaucoma. The mean age at onset of POAG was 27 years (23–38 years) in CYP1B1 mutation carriers compared with 51 years (48–64 years) in non-carriers (Vincent *et al.*, 2002). CYP1B1 is a member of the cytochrome P450 family which is thought to metabolize as yet unknown molecules critical for anterior chamber development (Stoilov *et al.*, 1997).

All of the above POAG loci were mapped or identified using single extended kindreds, in which the disease was consistent with autosomal dominant inheritance with high penetrance. In the great majority of cases, POAG is not a single gene disorder but a complex trait, where multiple loci are presumed to interact with the environment and with each other to influence the cause or progression of the disease. Although some preliminary genome-wide scans have been carried out in large numbers of POAG families, none has unambiguously mapped a disease-causing locus. The likely genetic heterogeneity adds an extra degree of difficulty in identifying susceptibility genes.

Primary angle closure glaucoma

Little is known about the molecular basis of the disease in this, the commonest, form of glaucoma in East or South East Asian populations. It results from permanent closure of the filtration angle as a result of iris apposition to the trabecular meshwork. It tends to occur in short hypermetropic (long-sighted) eyes with an anteriorly placed lens. The prevalence increases with age and in the presence of a family history and females are more often affected than males, but specific causal factors remain unknown.

Age-related macular degeneration

The most common cause of blindness in westernized countries is a disorder of the central retina called age-related macular degeneration (AMD). This accounts for about one-half of all blind registrations in Western countries and is increasingly recognised in less developed countries. It results from loss of the light-sensitive photoreceptors within a specialized region of the central retina known as the *macula* (Figure 32.1).

The macula is 5–6 mm in diameter and centred around the fovea, the region of maximal visual acuity, which contains the highest density of cone photoreceptors and 25% of all ganglion cells, the output cells from retina to brain (Hendrickson, 2005) (Figure 32.1). The importance of this region is emphasised by the fact that 40% of the primary visual cortex processes the central 1.5 mm diameter area of the fovea (Hendrickson, 2005). The macula also contains large numbers of rod photoreceptors, although these are absent from the central fovea (Figure 32.1). In the retina overall, rod photoreceptors outnumber cones by 20:1, and are exquisitely sensitive to light, but show poor spatial resolution.

AMD is the late stage of a condition called age-related maculopathy (ARM), which is present in about 5% of individuals at age 60 years, rising to about 30% at age 80 (Evans, 2003). The prevalence of AMD increases from 6% at ages 65–74 years to 11% at ages 75–84 years. ARM is characterized by both diffuse (basal) and focal (drusen) extracellular deposits within the macula. The reason for loss of (cone and rod) photoreceptor function within the macula is usually the presence of a long-standing disease process within the adjacent retinal pigment epithelium (RPE). RPE cells have an intimate metabolic, nutritional and trophic relationship with the overlying photoreceptors, including the supply and recycling of retinoids, phagocytosis of photoreceptor outer segments and release of neurotrophic factors required for maintenance of the choroidal vasculature and photoreceptors. The precise sequence of events remains uncertain but a plausible scenario would be that RPE cells and adjacent choroidal capillaries become increasingly compromised by age-associated oxidative damage, leading to exposure of modified self antigens, such as cross-linked or oxidized lipids and proteins, and activation of the alternative complement pathway. The resultant auto-immune attack leads to extracellular sub-RPE deposit formation and further changes within the RPE and choroidal vasculature, including growth of immature and fragile choroidal blood vessels through the thickened sub-RPE deposits, which leak and later bleed into the central retina (choroidal neovascularization, CNV). This results in a catastrophic loss of central vision (Penfold and Provis, 2005). This model is largely based on recent genetic findings that strongly implicate the alternative complement pathway, which is discussed below.

Both diffuse and focal deposits in ARM lie between a structure called Bruch's membrane and the RPE (Figure 32.1). Bruch's membrane is a pentalaminar basement membrane consisting of the basal lamina of the RPE, an inner collagenous layer, a central elastin layer, an outer collagenous layer and the basal lamina of the underlying choroidal capillaries. Blood flow through the choroidal capillaries nourishes the RPE, macula and outer retina (including the photoreceptor layer) and is one of the highest in the body (85% of ocular blood flow goes to the choroid compared with 4% to the retina).

The composition of the focal sub-RPE deposits, called drusen (literally ''bumps'') is known to include over 100 different proteins, as well as a variety of lipids such as cholesterol (Crabb *et al.*, 2002). However, one of the more abundant classes of proteins is a variety of complement components and complement regulatory proteins, including acute phase proteins, which form part of the innate immune system, the body's first line of defence against microbial invaders. The composition of the diffuse or basal sub-RPE deposits is less clear but again includes complement fragments and complement regulatory molecules as well as proteinaceous and lipid debris.

There are two major types of basal deposit. Firstly, deposits that lie between RPE and its basal lamina (called basal laminar deposits), which are different in composition from drusen, accumulate strongly with age, and have a less clear relationship to AMD-associated blindness than the second type, called basal linear deposits. The latter occur between the RPE basal lamina and inner collagenous layer of Bruch's membrane, resemble drusen in composition and are strongly associated with AMD.

There are a number of well established risk factors for AMD. The major risk factor is age. Western populations have an eight- to tenfold increased risk of AMD at age 90 years compared with age 50 years (Evans, 2003). The exponential increase in age-specific prevalence of both ARM and AMD is reminiscent of many age-associated disorders. A recent study of 897 elderly individuals in Iceland showed that 58% of those at least 80 years of age and all centenarians had either mild or severe forms of AMD (Geirsdottir *et al.*, 2005). This confirms the observation that the accumulation of both diffuse and focal sub-RPE deposits with age is nearly universal, and suggests that other risk factors merely accelerate or delay the inevitable. Other well-established risk factors include family history, smoking (odds ratio 2–3), raised levels of C-reactive protein, genetic variation at the *APOE* locus (odds ratio 0.5 for E4 allele) and ethnicity. Less well established risk factors include exposure to bright light, dietary and plasma levels of antioxidants, endogeneous and exogenous oestrogens (Evans, 2003).

The heritability of ARM and AMD has been investigated in two large twin studies (Hammond *et al.*, 2002; Seddon *et al.*, 2005). Both studies examined population based twin cohorts (UK, USA) and compared the prevalence of ARM or AMD in monozygotic ($N = 226, 210$) compared with dizygotic ($N = 280, 181$) twin pairs. In the US study, the age range of twins was 76–86 years, and the age-corrected estimates of heritability arising from additive effects varied from 0.53 to 0.73, depending on disease grading and comparisons, while the contribution of unique environmental factors was 0.20–0.33 (Seddon *et al.*, 2005). Shared environmental factors were not found to be statistically significant. In the second study, the concordance for ARM in MZ twins was 0.37, compared with 0.19 for DZ twins, and the heritability was 0.45 (Hammond *et al.*, 2002). Non-shared environment accounted for half (0.51) of the total variation in ARM.

Genetic influences in age-related macular degeneration

Extracellular matrix proteins

The fibulin family of secreted extracellular matrix (ECM) proteins consists of six members (fibulin1–6) which share a modular, elongated structure, including a central region with 5–11 calcium binding epidermal growth factor (cbEGF) domains and characteristic amino (N) and carboxyl (C) terminal domains (Timpl *et al.*, 2003). The importance of the high extracellular calcium concentration (~1 mM) to fibulins and other ECM proteins is emphasised by their changed structural and functional properties in its absence (Timpl *et al.*, 2003). The fibulins are thought to stabilise supramolecular ECM structures, such as basement membrane networks and elastic fibres, and to link such structures to cells. Two members of this family have been implicated in AMD or AMD-like disorders (Stone *et al.*, 1999; 2004). A heterozygous founder mutation in the fibulin-3 gene (*EFEMP1*), which changes a conserved arginine to a tryptophan residue at position 345 (Arg345Trp), was found in families with a rare autosomal dominant condition called Doyne honeycomb retinal dystrophy (DHRD, also called Malattia Leventinese). The retina in this condition shows multiple macular drusen, resembling a mosaic or honeycomb, similar to those seen in AMD, but causing severe visual loss in young adults. The *EFEMP1* gene is widely expressed but shows highest expression in the eye (especially

choroid/RPE), in certain capillaries and endothelial cells and in the lung (Stone *et al.*, 1999; Giltay *et al.*, 1999). The protein appears to be secreted from RPE cells although it does not normally localise to Bruch's membrane or choroid, or to the drusen found in DHRD (Marmorstein *et al.*, 2002). Curiously, while it is normally found in the interphotoreceptor matrix surrounding photoreceptors and at their synaptic terminals, in both DHRD and AMD retinas it is present as an abnormal extracellular deposit between RPE and Bruch's membrane, in regions where there are overlying drusen. The Arg345Trp mutant fibulin-3 was shown to be misfolded and to accumulate within RPE cells, rather than being secreted, suggesting that it may aggregate and compromise these cells. The fivefold lower rate of secretion of mutant protein may still be sufficient to form extracellular aggregates, which could reduce cell-matrix adhesion and become the target of immune attack and drusen formation (see below). Extracellular aggregates would be expected to form on the apical or retinal side of the RPE, rather than the Bruch's membrane or basal side, but either the mutant protein is not extruded by normal secretory mechanisms or RPE polarity is altered.

A variety of rare amino acid changing (missense) mutations have been found in evolutionarily conserved residues of the fibulin-5 gene (*FBLN5*) in 1–2% of AMD patients and no controls (Stone *et al.*, 2004; Lotery *et al.*, 2005). All of these *FBLN5* mutations, which span the entire protein, together with a single *Gln5346Arg* variant in the *FBLN6* gene, are associated both with a specific type of drusen, called cuticular or basal laminar drusen, and with RPE detachments, which are together present in about 20% of AMD subjects (Stone *et al.*, 2004). The *FBLN5* gene is expressed widely, including RPE cells and the elastic lamina of arteries, and its binding to both integrins and tropoelastin suggests that it connects extracellular elastin fibres to cells (Argraves *et al.*, 2003). Fibulin-5 has a single RGD (Arginine-Glycine-Aspartate) motif in the N-terminal domain which promotes cell adhesion by interaction with αvβ3, αvβ5 and α9β1 integrins (Nakamura *et al.*, 2002). Fibulin-5 also interacts both with fibrillin-1, another component of elastic fibres (Freeman *et al.*, 2005) and with extracellular superoxide dismutase (*SOD3*), which in fibulin-5 deficient mice results in an increase in vascular superoxide formation, which could promote aggregation of ECM proteins (Nguyen *et al.*, 2004).

Homozygous deficiency of fibulin-5 in mice and humans is associated with defective elastogenesis and the condition cutis laxa, with loose skin, diverticulosis, pulmonary emphysema and vascular abnormalities (Timpl *et al.*, 2003; Argraves *et al.*, 2003). In contrast, the mutations in the *FBLN5* gene associated with AMD were heterozygous and suggested a gain-of-function phenotype. Although the disease mechanism remains to be fully elucidated, it was suggested that, similar to DHRD, impaired RPE secretion of mutant proteins could lead to haploinsufficiency and impaired adhesion between elastin, a major component of Bruch's membrane, and RPE cells, either directly or indirectly via integrin-mediated cell attachment (Stone *et al.*, 2004).

Two other rare autosomal dominant forms of macular degeneration with sub-RPE deposits and clinical features resembling AMD result from mutations affecting the ECM proteins TIMP-3 and C1QTNF5 respectively (Weber *et al.*, 1994; Hayward *et al.*, 2003). Mutations in the tissue inhibitor of metalloproteinases 3 or *TIMP3* gene give rise to Sorsby fundus dystrophy, which is associated with sub-RPE deposits and macular degeneration resembling AMD. TIMP3 is involved in ECM remodeling and has been shown to inhibit matrix metalloproteinases, to bind fibulin-3 and to competitively inhibit binding of vascular endothelial growth factor (VEGF) to one of its receptors (VEGFR2 or KDR) (Qi *et al.*, 2003; Klenotic *et al.*, 2004). VEGF is a potent enhancer of angiogenesis and vascular permeability which is implicated in the choroidal neovascularisation of AMD. It remains unclear whether some or all of these mechanisms are operative in the sub-RPE deposit

formation or neovascularisation seen in Sorsby fundus dystrophy.

Similarly, the precise disease mechanism is unclear in late-onset retinal macular degeneration (L-ORMD), which closely resembles Sorsby fundus dystrophy and the common neovascular or ''wet'' form of AMD (Ayyagari *et al.*, 2000; Hayward *et al.*, 2003). This disorder is caused by a founder mutation in the short-chain collagen gene *C1QTNF5*, which leads to formation of high molecular weight aggregates and failure of C1QTNF5 protein secretion by RPE cells into the ECM (Hayward *et al.*, 2003; Shu *et al.*, 2006).

The innate immune system: complement factor H

Three studies reported a strong association between AMD and a polymorphism within the *CFH* gene encoding complement factor H (CFH) (Klein *et al.*, 2005; Edwards *et al.*, 2005; Haines *et al.*, 2005). In each report, an association was found between a common CFH amino acid substitution changing a conserved tyrosine to histidine residue (Tyr402His) within the seventh short consensus repeat (SCR) module of CFH. CFH is a modular protein with 20 SCR domains showing different binding specificities, which down-regulates the alternative complement activation pathway. A single copy of the *Tyr402His* allele was associated with a two- to three fold increased risk of AMD compared with unaffected controls, while two copies were associated with a 6–7 fold increased risk of AMD in case-control studies.

Two of the three studies identified the CFH variant in the course of following up a linkage signal on chromosome 1q31, which was found in several whole genome scans of affected sib pairs with AMD (Fisher *et al.*, 2005). These studies carried out case-control association studies using single nucleotide polymorphism (SNP) markers spanning the region (Haines *et al.*, 2005; Edwards *et al.*, 2005). The third study carried out a whole genome scan in 96 AMD cases and 50 controls using 103,611 SNPs, with an average density of 1 SNP every 30 kilobases (kb) (Klein *et al.*, 2005). In the latter study, a remarkable result was obtained in which only two SNPs showed statistically significant associations with AMD ($P < 4.8 \times 10^{-7}$) after using a conservative Bonferroni correction for multiple testing. One of these SNPs was excluded, leaving a single associated SNP located within an intron of the *CFH* gene. Both SNPs lay within a 500 kb block of strong linkage disequilibrium within the study sample, which is about tenfold larger than is typical for the human genome. This block contains the entire Regulator of Complement Activation (*RCA*) gene cluster, which includes seven complement regulatory genes, *CFH*, *FHL1* and *FHR1–5*. The authors subdivided this region further, using data from the normal population, and identified a high risk 6 SNP-haplotype spanning only the *CFH* gene, which conferred either a 4.6-fold or 7.4-fold increased risk of AMD in heterozygotes and homozygotes respectively. The entire *CFH* gene was then sequenced in 93 controls to identify all common variants. This identified 50 variant sites, three of which were predicted to change the amino acid sequence of the protein (non-synonymous substitutions), including the Tyr402His substitution within exon 9 (SCR module 7), which was present in 97% of the high risk haplotypes and showed the strongest association with AMD (Klein *et al.*, 2005). In an independent data set, the AMD risk for Tyr402His (CC) homozygotes, who represented 12% of this population, increased from about 20% in the general population to 54% (2.4-fold increase): in CT heterozygotes it was only marginally increased (22%), while in low risk (TT) homozygotes it was more than halved (9%) (data from Zareparsi *et al.*, 2005).

Strong supporting evidence also implicates the *CFH* gene in AMD. Firstly, CFH protein is present in choroidal blood vessels and in an area bordering RPE cells adjacent to Bruch's membrane (Klein *et al.*, 2005). Secondly, CFH is a key regulator of the alternative complement activation pathway, suggesting that risk alleles such as *402His* may be less effective than *402Tyr* in down-regulating the complement pathway, which could lead to

inappropriate activation and autoimmune damage to host tissues such as RPE and the choroidal vasculature. CFH is secreted in large amounts by the liver into the plasma which coats host cells and ECM surfaces. The alternative complement activation pathway normally shows continuous low level activation, by conversion of C3 to C3b by the C3 convertase (C3bBb). If C3b is formed in sufficient quantities, it leads to amplification of a destructive cascade with opsonization and phagocytosis of cells and lysis of target cells by formation of the terminal C5b-9 membrane attack complex (MAC) (Rodriguez de Cordoba *et al.*, 2004). This is therefore tightly controlled, both at the level of the cell membrane, by means of a group of membrane proteins (DAF, MCP, CR1, CD59), and by fluid phase regulators, particularly plasma CFH, which is continuously required to suppress C3b formation. CFH acts by promoting the decay of the C3 convertase (C3bBb) and by acting as a cofactor in the proteolytic inactivation of C3b by factor I. CFH is also synthesised outside the liver by tissues such as RPE and vascular endothelium, which appears to boost the local protection of tissues lacking membrane bound complement regulators.

The binding of CFH to host as opposed to activator (e.g. pathogen) cell surfaces is complex and involves binding of at least three SCR modules (SCR3, SCR13, SCR20) to host polyanionic heparin or sialic acid molecules, as well as other SCRs, using binding patterns that are unique to different activators (Pangburn *et al.*, 2000). Several common pathogens have developed receptors capable of binding CFH to escape detection, including *Streptococcus pyogenes* and *pneumoniae*, which bind SCR7, containing the *Tyr402His* variant (Rodriguez de Cordoba *et al.*, 2004). It is possible that this variant, which is predicted to alter the function of SCR7 (Rodriguez de Cordoba *et al.*, 2004), reduces the inhibitory effects of CFH on complement activation, perhaps as a result of natural selection to enhance complement lysis of such pathogens. A scenario in which both residues (402Tyr and 402His) have a selective advantage in the face of different microbial infections could lead to a balanced polymorphism, with maintenance of both alleles at high frequency. The penalty could be a less stringent inhibition of complement-mediated attack in later life, when synthesis of CFH by liver or RPE declines.

Homozygous loss-of-function mutations in the *CFH* gene lead to the rare condition type II membranoproliferative glomerulonephritis, which is associated with childhood onset of renal failure, electron-dense deposits under the renal glomerular basement membrane and, in 10% of patients, early-onset AMD with typical macular drusen and visual loss (Appel *et al.*, 2005). The glomerular podocytes lack membrane-bound complement regulators and so rely heavily on CFH to prevent complement activation. The glomerular capillary tuft-basal lamina-podocyte resembles the choroidal capillary–Bruch's membrane–RPE interface in its very high blood flow, the presence of fenestrated capillaries and the intimate contact between blood and the basal laminae of both RPE and podocytes (Appel *et al.*, 2005). This anatomical arrangement allows fast passage of fluids and large molecules (<70 kD in size, including protein-bound vitamin A), into and out of the retina, but perhaps contributes to the vulnerability of both sites to inflammatory attack and complement-mediated damage (Marshall *et al.*, 1998).

Many intriguing questions remain regarding the molecular pathology of AMD. One possible model emerging from the recent CFH findings is that, during ageing, RPE cells become increasingly detached from their basal lamina, which is aggravated by immune deposits. Bruch's membrane shows a linear increase in thickness with age, which is most marked in the macula, as a result of the build up of membrane, protein and lipid debris (Marshall *et al.*, 1998). Loss of cell adhesion can have wide ranging consequences, including altered signal transduction and changes in gene expression, differentiation, polarity, cell survival and apoptosis (Yamada and Geiger, 1997). Reduced expression of CFH and its shorter isoform, FHL1, both of which are synthesised by RPE cells (Hageman *et al.*, 2005) could follow, leading to

increased complement activation, drusen formation and a vicious cycle of sub-RPE deposit formation and complement-mediated damage. Why is the macula the preferred site of such damage? The density of metabolically active photoreceptors associated with each RPE cell is highest in this region, the choroidal blood flow and pO_2 are also higher than elsewhere in the retina and the central retina is most exposed to light and oxidative damage. In short, RPE cells in the macula may have the greatest exposure to a variety of oxidative, inflammatory, physical and other stresses throughout life, making them highly vulnerable to age-related damage.

The number of individuals over the age of 60 years is expected to triple from about 600 million in 2000 to 2 billion by 2050 so that, in the developed world, they are predicted to constitute about one-third of the population (United Nations, 2001). The burden of age-related blindness is therefore set to increase accordingly. This will impose a substantial burden on our health care systems, emphasising the need for urgency in developing improved methods of prevention and care. These are likely to be informed by the results of genetic studies of such conditions, which are already clarifying our understanding of disease pathogenesis.

Acknowledgements

I would like to thank Dr. Andrew Carothers for helpful discussions.

REFERENCES

Alward, W.L., Kwon, Y.H., Kawase, K. *et al.* (2003). Evaluation of optineurin sequence variations in 1,048 patients with open-angle glaucoma. *Am J Ophthalmol*, **136**, 904–10.

Appel, G.B., Cook, H.T., Hageman, G. *et al.* (2005). Membranoproliferative glomerulonephritis type II (dense deposit disease): an update. *J Am Soc Nephrol*, **16**, 1392–403.

Argraves, W.S., Greene, L.M., Cooley, M.A. and Gallagher, W.M. (2003). Fibulins: physiological and disease perspectives. *EMBO Rep*, **4**, 1127–31.

Aroca-Aguilar, J.D., Sanchez-Sanchez, F., Ghosh, S., Coca-Prados, M. and Escribano, J. (2005). Myocilin mutations causing glaucoma inhibit the intracellular endoproteolytic cleavage of myocilin between amino acids Arg226 and Ile227. *J Biol Chem*, **280**, 21043–51.

Ayyagari, R., Griesinger, I.B., Bingham, E. *et al.* (2000). Autosomal dominant hemorrhagic macular dystrophy not associated with the *TIMP3* gene. *Arch Ophthalmol*, **118**, 85–92.

Buhrmann, R.R., Quigley, H.A., Barron, Y. *et al.* (2000). Prevalence of glaucoma in a rural East African population. *Invest Ophthalmol Vis Sci*, **41**, 40–8.

Congdon, N.G. and Taylor, H.R. (2003). Age-related cataract. In Johnson, G.J., Minassian, D.C., Weale, R.A. and West, S.K. (eds.), *The epidemiology of eye disease*, pp.105–19. London: Arnold.

Conley, Y.P., Erturk, D., Keverline, A. *et al.* (2000). A juvenile-onset, progressive cataract locus on chromosome 3q21-q22 is associated with a missense mutation in the beaded filament structural protein-2. *Am J Hum Genet*, **66**, 1426–31.

Crabb, J.W., Miyagi, M., Gu, X. *et al.* (2002). Drusen proteome analysis: an approach to the etiology of age-related macular degeneration. *Proc Natl Acad Sci USA*, **99**, 14682–7.

Delaye, M. and Tardieu, A. (1983). Short-range order of crystallin proteins accounts for eye lens transparency. *Nature*, **302**, 415–17.

Edwards, A.O., Ritter, R., 3rd, Abel, K.J. *et al.* (2005). Complement factor H polymorphism and age-related macular degeneration. *Science*, **308**, 421–4.

Evans, J. (2003). Age-related macular degeneration. In Johnson, G.J., Minassian, D.C., Weale, R.A. and West, S.K. (eds.), *The epidemiology of eye disease*, pp. 356–69. London: Arnold.

Fisher, S.A., Abecasis, G.R., Yashar, B.M. *et al.* (2005). Meta-analysis of genome scans of age-related macular degeneration. *Hum Mol Genet*, **14**, 2257–64.

Francis, P., Berry, V., Bhattacharya, S. and Moore, A. (2000). Congenital progressive polymorphic cataract caused by a mutation in the major intrinsic protein of the lens, MIP (AQP0). *Br J Ophthalmol*, **84**, 1376–9.

Freeman, L.J., Lomas, A., Hodson, N. *et al.* (2005). Fibulin-5 interacts with fibrillin-1 molecules and microfibrils. *Biochem J*, **388**, 1–5.

Geirsdottir, A., Stefansson, E., Jonasson, F. *et al.* (2005). Do all individuals with a family history of age-related maculopathy (ARM) develop age-related macular degeneration (AMD) if they live to be 100 years old? *Invest Ophthalmol Vis Sci*, A3310.

Giltay, R., Timpl, R. and Kostka, G. (1999). Sequence, recombinant expression and tissue localization of two novel extracellular matrix proteins, fibulin-3 and fibulin-4. *Matrix Biol*, **18**, 469–80.

Gong, G., Kosoko-Lasaki, O., Haynatzki, G. R. and Wilson, M. R. (2004). Genetic dissection of myocilin glaucoma. *Hum Mol Genet*, **13** Spec No 1, R91–102.

Graw, J. (2004). Congenital hereditary cataracts. *Int J Dev Biol*, **48**, 1031–44.

Hageman, G. S., Anderson, D. H., Johnson, L. V. *et al.* (2005). A common haplotype in the complement regulatory gene factor H (*HF1/CFH*) predisposes individuals to age-related macular degeneration. *Proc Natl Acad Sci USA*, **102**, 7227–32.

Haines, J. L., Hauser, M. A., Schmidt, S. *et al.* (2005). Complement factor H variant increases the risk of age-related macular degeneration. *Science*, **308**, 419–21.

Hammond, C. J., Duncan, D. D., Snieder, H. *et al.* (2001). The heritability of age-related cortical cataract: the twin eye study. *Invest Ophthalmol Vis Sci*, **42**, 601–5.

Hammond, C. J., Snieder, H., Spector, T. D. and Gilbert, C. E. (2000). Genetic and environmental factors in age-related nuclear cataracts in monozygotic and dizygotic twins. *N Engl J Med*, **342**, 1786–90.

Hammond, C. J., Webster, A. R., Snieder, H. *et al.* (2002). Genetic influence on early age-related maculopathy: a twin study. *Ophthalmology*, **109**, 730–6.

Harding, J. J. (1991). *Cataract: biochemistry, epidemiology and pharmacology*. London: Chapman and Hall.

Hayward, C., Shu, X., Cideciyan, A. V. *et al.* (2003). Mutation in a short-chain collagen gene, CTRP5, results in extracellular deposit formation in late-onset retinal degeneration: a genetic model for age-related macular degeneration. *Hum Mol Genet*, **12**, 2657–67.

Heijl, A., Leske, M. C., Bengtsson, B. *et al.* (2002). Early Manifest Glaucoma Trial Group. Reduction of intraocular pressure and glaucoma progression: results from the Early Manifest Glaucoma Trial. *Arch Ophthalmol*, **120**, 1268–79.

Hejtmancik, J. F. and Kantorow, M. (2004). Molecular genetics of age-related cataract. *Exp Eye Res*, **79**, 3–9.

Hejtmancik, J. F., Kaiser-Kupfer, M. I. and Piatigorsky, J. (2001). Molecular biology and inherited disorders of the eye lens. In Scriver, C. R. and Sly, W. S. (eds.), *The metabolic basis of inherited disease*, 8th edn, pp. 6033–62. New York: McGraw Hill.

Hendrickson, A. (2005). Organization of the adult primate fovea. In Penfold, P. L. and Provis, J. M. (eds.), *Macular degeneration*, pp.1–23. Berlin: Springer-Verlag.

Iyengar, S. K., Klein, B. E., Klein, R. *et al.* (2004). Identification of a major locus for age-related cortical cataract on chromosome 6p12-q12 in the Beaver Dam Eye Study. *Proc Natl Acad Sci USA*, **101**, 14485–90.

Johnson, L. V. and Anderson, D. H. (2004). Age-related macular degeneration and the extracellular matrix. *N Engl J Med*, **351**, 320–2.

Johnstone, G. J. and Quigley, H. A. (2003). In Johnson, G. J., Minassian, D. C., Weale, R. A. and West, S. K. (eds.), *The epidemiology of eye disease*, pp. 222–39. London: Arnold.

Klein, R. J., Zeiss, C., Chew, E. Y. *et al.* (2005). Complement factor H polymorphism in age-related macular degeneration. *Science*, **308**, 385–9.

Klenotic, P. A., Munier, F. L., Marmorstein, L. Y. and Anand-Apte, B. (2004). Tissue inhibitor of metalloproteinases-3 (TIMP-3) is a binding partner of epithelial growth factor-containing fibulin-like extracellular matrix protein 1 (EFEMP1). Implications for macular degenerations. *J Biol Chem*, **279**, 30469–73.

Lee, S. C., Wang, Y., Ko, G. T. *et al.* (2001). Risk factors for cataract in Chinese patients with type 2 diabetes: evidence for the influence of the aldose reductase gene. *Clin Genet*, **59**, 356–9.

Lotery, A. J., Avery, K., Goverdhan, S. V. *et al.* (2005). Mutation analysis of fibulin 5 in a United Kingdom cohort of age-related macular degeneration patients. *Invest Ophthalmol Vis Sci*, A1150.

Marmorstein, L. Y., Munier, F. L., Arsenijevic, Y. *et al.* (2002). Aberrant accumulation of, EFEMP1 underlies drusen formation in Malattia Leventinese and age-related macular degeneration. *Proc Natl Acad Sci USA*, **99**, 13067–72.

Marshall, J., Hussain, A. A., Starita, C., Moore, D. J. and Patmore, A. L. (1998). Aging and Bruch's membrane. In Marmor, M. F. and Wolfensberger, T. J. (eds.), *The retinal pigment epithelium*, pp. 669–92. New York: Oxford University Press.

Martin, N., Boomsma, D. and Machin, G. (1997). A twin-pronged attack on complex traits. *Nat Genet*, **17**, 387–92.

McIlwaine, J. T. (1996). *An introduction to the biology of vision*. Cambridge, MA: Cambridge University Press, pp. 3–5.

Nakamura, T., Lozano, P. R., Ikeda, Y. *et al.* (2002). Fibulin-5/DANCE is essential for elastogenesis in vivo. *Nature*, **415**, 171–5.

Nguyen, A. D., Itoh, S., Jeney, V. *et al.* (2004). Fibulin-5 is a novel binding protein for extracellular superoxide dismutase. *Circ Res*, **95**, 1067–74.

Nguyen, T. D., Chen, P., Huang, W. D. *et al.* (1998). Gene structure and properties of TIGR, an olfactomedin-related glycoprotein cloned from glucocorticoid-induced trabecular meshwork cells. *J Biol Chem*, **273**, 6341–50.

Okano, Y., Asada, M., Fujimoto, A. *et al.* (2001). A genetic factor for age-related cataract: identification and characterization of a novel galactokinase variant, ''Osaka,'' in Asians. *Am J Hum Genet*, **68**, 1036–42.

Pangburn, M. K., Pangburn, K. L., Koistinen, V., Meri, S. and Sharma, A. K. (2000). Molecular mechanisms of target recognition in an innate immune system: interactions among factor H, C3b, and target in the alternative pathway of human complement. *J Immunol*, **164**, 4742–51.

Penfold, P. L. and Provis, J. M. (Eds.) (2005). *Macular degeneration*. Heidelberg: Springer-Verlag.

Qi, J. H., Ebrahem, Q., Moore, N. *et al.* (2003). A novel function for tissue inhibitor of metalloproteinases-3 (TIMP3): inhibition of angiogenesis by blockage of VEGF binding to VEGF receptor-2. *Nat Med*, **9**, 407–15.

Quigley, H. A. (2004). New paradigms in the mechanisms and management of glaucoma. *Eye*, Advance Online Publication 29 October, 2004.

Reddy, M. A., Francis, P. J., Berry, V., Bhattacharya, S. S. and Moore, A. T. (2004). Molecular genetic basis of inherited cataract and associated phenotypes. *Surv Ophthalmol*, **49**, 300–15.

Reidy, A., Minassian, D. C., Vafidis, G. *et al.* (1998). Prevalence of serious eye disease and visual impairment in a north London population: population based, cross sectional study. *BMJ*, **316**, 1643–6.

Ren, Z., Li, A., Shastry, B. S. *et al.* (2000). A 5-base insertion in the gammaC-crystallin gene is associated with autosomal dominant variable zonular pulverulent cataract. *Hum Genet*, **106**, 531–7.

Resnikoff, S., Pascolini, D., Etya'ale, D. *et al.* (2004). Global data on visual impairment in the year 2002. *Bull World Health Organ*, **82**, 844–51.

Rezaie, T., Child, A., Hitchings, R. *et al.* (2002). Adult-onset primary open-angle glaucoma caused by mutations in optineurin. *Science*, **295**, 1077–9.

Rodriguez de Cordoba, S., Esparza-Gordillo, J., Goicoechea de Jorge, E., Lopez-Trascasa, M. and Sanchez-Corral, P. (2004). The human complement factor H: functional roles, genetic variations and disease associations. *Mol Immunol*, **41**, 355–67.

Seddon, J. M., Cote, J., Page, W. F., Aggen, S. H. and Neale, M. C. (2005). The US twin study of age-related macular degeneration: relative roles of genetic and environmental influences. *Arch Ophthalmol*, **123**, 321–7.

Shu, X., Tulloch, B., Lennon, A. *et al.* (2006). Disease mechanisms in late-onset retinal macular degeneration associated with mutation in *C1QTNF5*. *Hum Mol Genet*, **15**, 1680–9.

Stoilov, I., Akarsu, A. N. and Sarfarazi, M. (1997). Identification of three different truncating mutations in cytochrome P4501B1 (*CYP1B1*) as the principal cause of primary congenital glaucoma (Buphthalmos) in families linked to the *GLC3A* locus on chromosome 2p21. *Hum Mol Genet*, **6**, 641–7.

Stone, E. M., Braun, T. A., Russell, S. R. *et al.* (2004). Missense variations in the fibulin 5 gene and age-related macular degeneration. *N Engl J Med*, **351**, 346–53.

Stone, E. M., Fingert, J. H., Alward, W. L. *et al.* (1997). Identification of a gene that causes primary open angle glaucoma. *Science*, **275**, 668–70.

Stone, E. M., Lotery, A. J., Munier, F. L. *et al.* (1999). A single EFEMP1 mutation associated with both Malattia Leventinese and Doyne honeycomb retinal dystrophy. *Nat Genet*, **22**, 199–202.

Taylor, H. R., West, S. K., Rosenthal, F. S. *et al.* (1998). Effect of ultraviolet radiation on cataract formation. *N Engl J Med*, **319**, 1429–33.

Tielsch, J. M., Sommer, A., Katz, J. *et al.* (1991). Racial variations in the prevalence of primary open-angle glaucoma. The Baltimore Eye Survey. *JAMA*, **266**, 369–74.

Timpl, R., Sasaki, T., Kostka, G. and Chu, M. L. (2003). Fibulins: a versatile family of extracellular matrix proteins. *Nat Rev Mol Cell Biol*, **4**, 479–89.

United Nations (2001). World population prospects. The 2000 revision highlights. www.un.org/esa/population/publications/wpp2000/highlights.pdf

Vincent, A. L., Billingsley, G., Buys, Y. *et al.* (2002). Digenic inheritance of early-onset glaucoma: *CYP1B1*, a potential modifier gene. *Am J Hum Genet*, **70**, 448–60.

Weber, B. H., Vogt, G., Pruett, R. C., Stohr, H. and Felbor, U. (1994). Mutations in the tissue inhibitor of metalloproteinases-3 (TIMP3) in patients with Sorsby's fundus dystrophy. *Nat Genet*, **8**, 352–6.

West, S. K., Munoz, B., Schein, O. D., Duncan, D. D. and Rubin, G. S. (1998). Racial differences in lens opacities: the Salisbury Eye Evaluation (SEE) project. *Am J Epidemiol*, **148**, 1033–9.

Wiggs, J. L., Auguste, J., Allingham, R. R. *et al.* (2003). Lack of association of mutations in optineurin with disease in patients with adult-onset primary open-angle glaucoma. *Arch Ophthalmol*, **121**, 1181–3.

Williams-Lyn, D., Flanagan, J., Buys, Y. *et al.* (2000). The genetic aspects of adult-onset glaucoma: a perspective from the Greater Toronto area. *Can J Ophthalmol*, **35**, 12–7.

Yamada, K. M. and Geiger, B. (1997). Molecular interactions in cell adhesion complexes. *Curr Opin Cell Biol*, **9**, 76–85.

Young, T. (1802). On the theory of light and colours. *Philos Trans R Soc London*, **92**, 12–48.

Zareparsi, S., Branham, K. E., Li, M. *et al.* (2005). Strong Association of the Y402H Variant in Complement Factor H at 1q32 with Susceptibility to Age-Related Macular Degeneration. *Am J Hum Genet*, **77**, 149–53.

33

Genetic and environmental influences on hearing impairment

Karen P. Steel

Prevalence of hearing impairment

Hearing impairment is undoubtedly a common disease. Around 1.06 per 1000 children are born with a significant, permanent hearing impairment (40 dB or greater increase in threshold in their better hearing ear), and by the age of nine years, this number has risen to around 1.65 per 1000 (Fortnum *et al.*, 2001). The prevalence of hearing impairment continues to increase with each decade of life, until 40% of the 71–80 years age group and 80% of the 80+ age group have a hearing loss of 35 dB or more (Davis, 1989; Davis and Moorjani, 2002). In total, approximately 20% of all adults over 18 in the UK suffer some form of hearing impairment (25 dB or greater hearing loss in at least one ear), and the proportions for other countries are very similar (Davis and Moorjani, 2002). The increase at various impairment levels is illustrated in Figure 33.1. However, thresholds are a crude reflection of the impairment, because it is not just the amplitude but the clarity of hearing that is affected. Our ability to distinguish speech sounds, to focus on specific sound sources such as one speaker in a noisy room, and to localize sounds in the environment all require accurate frequency and temporal discrimination, features that are disproportionately affected by hearing impairment. Much hearing loss with age affects sensitivity to high frequencies first, although some types of hearing loss have a more even effect across the frequency spectrum. Figure 33.3 illustrates typical hearing levels in age-related progressive hearing loss (Rosenhall, 2002).

Mechanisms of hearing impairment

Hearing depends upon the middle ear for collecting, amplifying and delivering the vibration of sound to the highly specialised cochlea, where sensory hair cells in the spiral organ of Corti detect these mechanical signals and convert them to electrical changes within the cell (auditory transduction). Depolarization of the cell allows synaptic release, triggering action potentials in cochlear neurons. Hair cells have an array of stereocilia (modified microvilli) at their upper surface, with extracellular

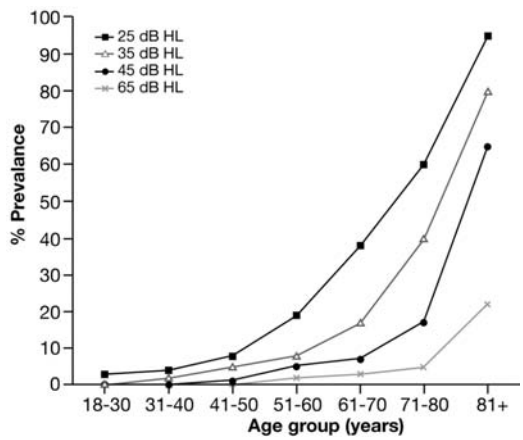

Figure 33.1 Prevalence of hearing loss (HL) with age, for different severities in decibels (db).

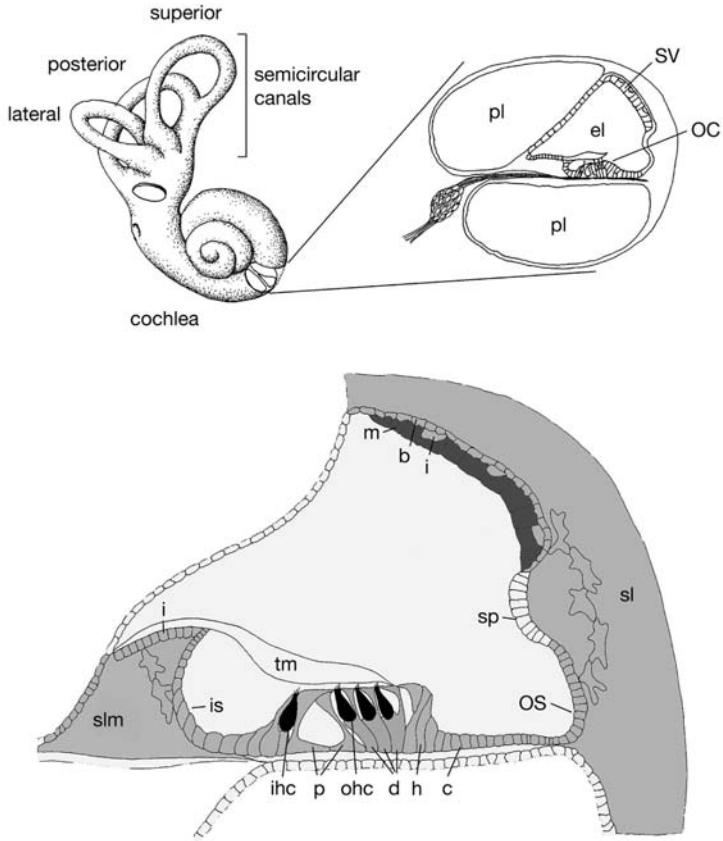

Figure 33.2 Anatomy of the cochlea. Cross section through the cochlear canal showing the organ of Corti (OC) on the floor of the endolymph (el) filled cochlear duct with the stria vascularis (SV) on the lateral wall and perilymph (pl) filled scala vestibuli and scala tympani (upper figures). Cross section of cochlear duct (scala media) showing the tectorial membrane (tm), inner hair cells (ihc), pillar cells (p); outer hair cells (ohc), Dieter's (d), Hensen's (h) and Claudius' (c) cells, outer sulcus (os), inner sulcus (is), spiral ligament (sl), spiral limbus (slm), interdental cells (i), marginal (m), basal (b) and intermediate (i) cells (lower figure).

links between them. When the stereocilia (or hair) bundle is deflected by incoming vibration, one type of link, the tip link, mechanically opens the transduction channel allowing rapid influx of cations into the cell. Hair cell function depends upon precise control of the environment, and specialized supporting cells within the organ of Corti and cellular structures around the cochlear duct serve this function. For example, the stria vascularis on the lateral wall pumps out potassium into the fluid bathing the tops of the hair cells, generating a high potassium concentration and a high resting potential, the endocochlear potential, which are both essential for hair cell function (Figure 33.2).

It is frequently stated that age-related hearing loss is due to degeneration of the sensory hair cells in the organ of Corti of the cochlea. However, there is no evidence for this contention from animal studies. Rather, it appears that hair cell degeneration is a correlate or a consequence of some primary dysfunction, either of hair cells or of some other part of the auditory system.

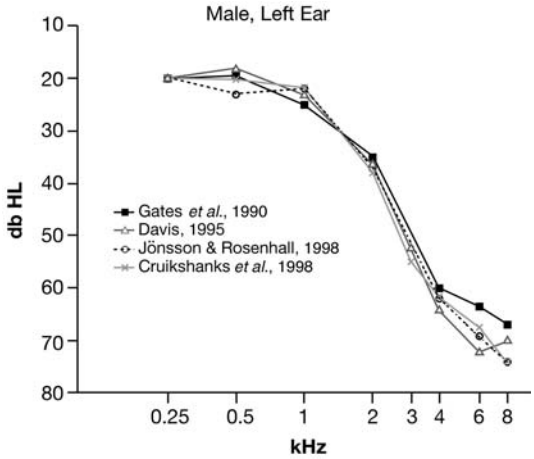

Figure 33.3 Typical audiograms of 70–80 years old, sensory type. Hearing loss (HL) is shown in decibels (db).

These sensory hair cells seem to be particularly sensitive to any disturbance of their homeostasis, resulting in their degeneration. However, findings in animal models suggest that hearing impairment correlates with hair cell function rather than with hair cell death, as there can be plenty of surviving but dysfunctional hair cells in a mouse with a *Tmc1* or a *Cdh23* mutation, or in a mouse with a mutation affecting cochlear homeostasis. Even with noise-induced damage, cochlear responses correlate better with stereocilia bundle damage than with hair cell death (e.g. Holme and Steel, 2004).

Mature mammalian hair cells in the cochlea never regenerate, so hair cell loss accumulates with age and this can be observed in human temporal bone specimens from people with hearing impairment. However, donated human inner ears generally represent the end-stage of a long disease process. Schuknecht and Gacek (1993) proposed several different fundamental types of hearing loss based on observations of many temporal bones and audiograms of people with progressive hearing loss with age (also known as presbyacusis). The most common type was sensory presbyacusis, characterised by the steeply sloping audiogram with poor thresholds at high frequencies, as illustrated in Figure 33.3, with associated hair cell loss at the basal end of the cochlear duct (the region normally involved in high frequency detection). The second type was strial presbyacusis, with more equal hearing loss across the frequency range and atrophy of the stria vascularis, a structure on the lateral wall of the cochlear duct that pumps high levels of potassium into the fluid bathing the upper surface of the hair cells and generates a high resting potential in this fluid. A third type, neural presbyacusis, was rare but showed loss of cochlear neurons leading to limited threshold increases but difficulty in speech discrimination. This is a useful first step in grouping pathological mechanisms. However, as we discover more of the genes underlying hearing impairment (see later), and understand more about their involvement in the pathological process, it is becoming clear that there are dozens, maybe hundreds, of different ways that hearing can be compromised. Furthermore, other parts of the auditory system can be involved in progressive hearing loss, such as the middle ear (Browning and Gatehouse, 1992; Rosowski *et al.*, 2003). Cognitive decline with age may interact with hearing impairment to exacerbate the functional disability in the elderly population, but the vast majority of cases of hearing loss are the result of a pathological process affecting the inner or middle ear (mainly the inner ear) rather than an effect of the central auditory system alone.

Treatments

There are no treatments available, only two main types of prosthesis. First is the hearing aid, which amplifies incoming sounds and can be adjusted to amplify certain frequencies more than others to match the pattern of hearing loss of an individual, and to match the limited dynamic range of the damaged ear to avoid painfully loud sounds being delivered. However, there is much to be learnt about how we use the temporal and frequency cues in speech and other sounds in order to improve programming of hearing aids to facilitate use of

these cues. The threshold for provision of a hearing aid is often considered to be around 25 dB hearing level, although many people are not fitted with an aid until their hearing is much worse than this (Davis and Moorjani, 2002). Second is the cochlear implant, which involves surgery to place an extended array of electrodes within the cochlear duct and a subcutaneous receiver for detecting the coded stimulation and transmitting to the electrode array. This is normally considered only for people who have such a profound hearing impairment that they are not helped by hearing aids. For example, the vast majority of children fitted with a cochlear implant have hearing loss of 95 dB or more (Fortnum *et al.*, 2002). Both prosthetic approaches require time to adjust, and in the case of cochlear implants, a considerable period of rehabilitation is needed to maximise benefit.

Neither prosthetic approach is ideal, so there is a need for other approaches to minimise the effects of hearing impairment. Understanding the molecular basis of hearing impairment should allow the development of treatments to stop or slow down progression of hearing loss, if not to restore hearing. Some preliminary work in animal models suggests that biological agents may be useful, at least if applied at around the time of a damaging stimulus (e.g. Wang *et al.*, 2003; Zhai *et al.*, 2004). Treatments might involve some way of triggering regeneration of lost hair cells and their supporting cells, because these cells do not regenerate once they die. Alternatively, a drug-based approach might allow some degree of transcriptional control, such as upregulating alternative genes to replace dysfunctional elements of the hearing process (e.g. Steel, 2000). Slowing down or stopping the advance of a progressive hearing loss seems to be a much more tractable biological problem than developing a treatment to correct an early developmental defect, and most deafness in the human population is progressive, even during early childhood (Fortnum *et al.*, 2001; Johansen *et al.*, 2004). The limited options currently available for hearing-impaired people together with the large numbers of affected individuals argues for further research towards understanding the molecular basis of the disease process.

Environmental causes of deafness

It is well established that exposure to certain environmental factors will lead to hearing loss (e.g. Fransen *et al.*, 2003). Excessive noise exposure is the most obvious cause of hearing loss. However, infections such as prenatal rubella or meningitis can lead to deafness. Certain drugs also have ototoxic effects, including aminoglycoside antibiotics, cisplatin and diuretic agents. In some cases it is clear that one of these factors is the immediate cause of deafness, such as hearing impairment that develops over the few days after exposure to an ototoxic drug or after meningitis, but often it is not so obvious. For example, although many studies of noise exposure in animals have shown the deleterious effects of noise on auditory function, such experiments are impossible in humans so we are left trying to piece together a history of noise exposure long after the event. The National Study of Hearing in the UK found a small but significant effect of relatively high reported noise exposure on hearing levels, but not for moderate noise exposures, and the effects were very small compared with the effects of age (Davis, 1989).

Animal studies have revealed some factors that moderate or potentiate noise-induced hearing loss. For example, exposure to certain drugs and chemicals like toluene and styrene or concomitant hyperthermia potentiate noise-induced hearing loss (Davis *et al.*, 2002; Fechter, 2004). Surprisingly, exposure to a stressful stimulus, such as heat stress, restraint stress or moderate noise levels, during the few days prior to the damaging noise exposure can reduce the amount of noise-induced damage (e.g. Yoshida *et al.*, 1999; Yoshida and Liberman, 2000). This phenomenon is known as conditioning or toughening.

Role of genetics in deafness in the population

Many reports of clinical populations of children with permanent hearing impairment have been published, usually suggesting between a third and a half of children have a genetic etiology (e.g. Fortnum *et al.*, 2002; Morzaria *et al.*, 2004). However, the criteria for ascribing a genetic aetiology often are not clear: many authors assume a genetic causation in the absence of any other clear cause, while others report syndromic deafness as a separate category even when it is highly likely to be due to a single gene mutation. Other reports of the role of single gene mutations in hearing impairment use a selected clinical population as the denominator. Examples of this are the reports of the frequent involvement of mutations of the *GJB2* gene in deafness (in some Mediterranean populations up to 50%): the clinical populations reported often include only those with a definitive family history suggesting autosomal recessive inheritance of severe or profound deafness, excluding the sporadic cases where only one child is affected and excluding those with mild or moderate hearing impairment. Nonetheless, all authors agree that genetics has an important role in childhood deafness.

What of age-related hearing loss? As mentioned above, hearing impairment increases with age in the population, and age of onset for an individual can be at any age (Figure 33.1). Three studies suggest that genetics plays an important role in this later-onset hearing loss. The first report looked at intraclass correlations in male twins from a Swedish register (Karlsson *et al.*, 1997). A total of 250 monozygotic pairs and 307 dizygotic pairs aged between 36 and 80 years old were examined by audiometry. Analysis of audiograms suggested that the hearing levels of the subjects were representative of the Swedish male population as a whole, and thresholds for high frequencies (3, 4, 6 and 8 kHz) were grouped together for study. The intraclass correlations in monozygotic twins decreased from 0.716 in the youngest group to 0.516 in the oldest group, while for dizygotic twins the correlations were much smaller. This suggests that genetic effects play a very significant role in explaining the variance observed, and also that environmental effects play an increasing role with increasing age. Heritability for the younger age groups was very high, approaching one.

A second study used the Framingham, Massachusetts, cohort which has been investigated at regular intervals over many years for a number of traits (Gates *et al.*, 1999). After excluding cases with known causes of hearing impairment such as trauma, surgery or Ménière disease, and cases with unilateral impairment or evidence of middle ear disease (air-bone gap >15 dB), 1079 members of the original cohort remained to be analyzed together with 1232 of their offspring. Thresholds for low, middle and high frequencies were measured by audiometry and were adjusted for age and sex using the population standards. The authors asked if adjusted thresholds in related pairs were more highly correlated than in unrelated pairs. The related pairs included sibs and parent-child pairs. Strong familial aggregations of age-adjusted thresholds were found, especially for mother–daughter and sister pairs, with heritability calculated to be between 0.26 and 0.35, and high frequency thresholds showed the greatest correlation in the overall analysis. When sensorineural hearing loss was analysed separately (assessed by the sloping shape of the audiogram towards high frequencies, as illustrated in Figure 33.3) low and mid-frequencies showed the largest heritability scores, up to 0.55. A later analysis of this cohort used the largest 328 families comprising 1789 individuals to carry out linkage analysis (DeStefano *et al.*, 2003). Six chromosomal regions showed log likelihood (LOD) scores of 1.5 or more for low or mid-frequency thresholds: 3 regions of chromosome 11, plus regions on chromosomes 10, 14 and 18. The implicated regions on 11q13.5, 11p and 14q include genes known to be involved in Usher syndrome type 1 (severe or profound childhood deafness, vestibular dysfunction and progressive retinitis

pigmentosa): Usher 1B caused by *MYO7A* mutations; Usher 1C caused by harmonin mutations; and Usher 1A, which is not yet identified, respectively. *MYO7A* is also involved in non-syndromic deafness DFNB2 and DFNA11. The 11q25 location implicated in age-related hearing loss overlaps with DFNB20, another non-syndromic deafness locus for which we have not yet found the responsible gene. These genes represent good candidates for searching for mutations associated with later-onset progressive hearing loss.

A third study used self-reported hearing loss by questionnaire in twins aged 70 years and older from the Danish twin registry (Christensen *et al.*, 2001). This analysis suggested a heritability of 0.4, which the authors considered to be an underestimate because of the limitations of self-report compared with audiometry.

Genes and environment

Some cases of hearing impairment result primarily from a single gene mutation inherited in a Mendelian fashion, and other cases are the consequence of a known environmental insult such as meningitis or cisplatin administration. However, in many (maybe most) cases, both the genome of an individual and the environment interact to lead to deafness. For example, not all cases of meningitis lead to hearing loss, and people working in the same noisy environment do not all develop the same degree of hearing impairment, and the genetic susceptibility of the individuals probably accounts for these differences. We already know one mutation, A1555G in the mitochondrial genome, which makes carriers especially sensitive to aminoglycoside-induced hearing loss. This same mutation is associated with age-related progressive hearing loss in some families, irrespective of exposure to any drug. Gene variants that make carriers more susceptible to noise-induced deafness are difficult to find in humans, because of the difficulty in determining the exact level of historical exposure, but we do know of a number of such variants from work with animal models. These genes include Cu/Zn-superoxide dismutase (*Sod1*) and glutathione peroxidase (*Gpx1*), which are both involved in reducing reactive oxygen species. Mice with either of these genes knocked out showed increased susceptibility to noise-induced hearing loss (Ohlemiller *et al.*, 1999; 2000). Heterozygotes for the *Sod1* knockout showed intermediate levels of the enzyme in blood and intermediate susceptibility to noise damage, but overexpression of *Sod1* did not lead to extra protection from damage (Ohlemiller *et al.*, 1999; Coling *et al.*, 2003). Mutations in two genes known to be involved in function of the stereocilia of sensory hair cells are reported to enhance noise sensitivity in heterozygotes: cadherin 23 (*Cdh23*), which has been proposed as a component of the extracellular tip links found between adjacent stereocilia of the hair bundle, and the plasma membrane calcium ATPase isoform 2 (*Atp2b2*), which encodes a calcium pump (Holme and Steel, 2004; Kozel *et al.*, 2002). Vasodilator-stimulated phosphoprotein (*Vasp*) which is expressed in pillar cells of the organ of Corti and is thought to be involved in cytoskeletal dynamics also appears to have a protective effect because mice with no functional *Vasp* gene have increased sensitivity to noise-induced damage (Schick *et al.*, 2004). Mice with an inactivated heat shock factor 1 gene (*Hsf1*) suffer increased degeneration of sensory hair cells in the cochlea (Sugahara *et al.*, 2003). Three genes involved in neuronal function have a protective role, as mutants have increased sensitivity to noise: the glutamate transporter GLAST encoded by *Slc1a3*, the nociceptin/orphaninFQ receptor encoded by *Oprl1*, and the α9 nicotinic acetylcholine receptor encoded by *Chrna9* (Hakuba *et al.*, 2000; Nishi *et al.*, 1997; Maison *et al.*, 2002). In contrast, mice with no acid-sensing ion channel 2 (*Accn1/Asic2*) are more resistant to temporary (but not permanent) threshold shifts following noise exposure (Peng *et al.*, 2004). All of these genes represent good candidates for involvement in human sensitivity to noise-induced damage to hearing.

Despite the observation that susceptibility to noise-induced hearing loss often parallels age-related hearing loss in mice (Holme and Steel, 2004; Vázquez *et al.*, 2004), this is not always the case (Yoshida *et al.*, 2000), suggesting that some specific pathological mechanisms may be involved as well as common pathways to hearing loss.

Genetic heterogeneity

It is clear that deafness is an extremely heterogeneous disease. There are likely to be hundreds of different genes involved, any one of which can underlie deafness. For syndromic deafness, there are over 400 distinct Mendelian disorders that include deafness as one of the features listed in Online Mendelian Inheritance in Man (OMIM). For non-syndromic deafness, over 100 loci have been found and 36 of the genes have been identified (Van Camp and Smith, 2005; Petit *et al.*, 2001; Bitner-Glindzicz, 2002; Friedman and Griffith, 2003). The genes represent a wide variety of molecules, ranging from myosin motors to transcription factors, ion channels to extracellular matrix components. They are expressed in diverse cell types within the auditory system, but mostly within the inner ear (Van Camp and Smith, 2005, see link to expression). As the genes involved in non-syndromic deafness have been localized, an interesting feature has emerged: recessive deafness is mostly early-onset childhood deafness, while the dominantly inherited forms are more often later-onset progressive hearing losses (Van Camp *et al.*, 1997). The age of onset can vary within and between families, with some cases showing onset in their forties or even later. Furthermore, a number of the genes are involved in both dominant and recessive deafness in different families. *MYO7A* is a good example of the complexity: it is involved in Usher syndrome type 1B (severe or profound childhood deafness, vestibular dysfunction and progressive retinitis pigmentosa), atypical Usher syndrome with progressive hearing loss and highly variable retinitis pigmentosa, dominant non-syndromic progressive deafness (DFNA11) and recessive non-syndromic deafness (DFNB2). Furthermore, there are at least seven different genes underlying Usher syndrome type 1 (Van Camp and Smith, 2005). This complexity can sometimes be explained by the molecular lesion involved, but not always. Variability in expression and penetrance can be a feature even within a family carrying the same mutation, and this is particularly the case for dominantly inherited deafness.

Heterogeneity is a feature of mouse too, as there are over a hundred different genes identified that are associated with some sort of developmental or functional defect of the auditory system (Anagnostopoulos, 2002). The earliest mouse deafness genes to be found were associated with overt balance defects, because these lead to characteristic head-bobbing and circling behavior, whereas deafness alone is not normally noticed. More recently phenotype-driven screening programs worldwide have specifically sought mutants with hearing impairment alone. Deafness in mice with targeted mutations is more often discovered due to heightened awareness of this phenotype. Hence, we now know of a variety of single genes involved in progressive hearing loss without balance defects. This broadens the range of candidate genes for investigation in human populations with age-related hearing loss. Nonetheless, single genes discovered because of a balance defect are still proving to be very important in understanding human deafness, for several reasons. Milder mutations may lead to hearing loss alone (for example, different *CDH23* mutations may cause profound deafness and balance defects or hearing loss alone). Many cases of deafness in childhood show vestibular dysfunction too, although this is often not reflected in overt balance problems (Huygen and Verhagen, 1994). Modifiers may influence the effects of these primary mutations on balance. Thus, any genes found to be involved in deafness in mice should represent good candidates for human deafness whether early or late-onset.

Modifiers of deafness genes

Several modifier genes have been localised that modify the phenotype of carriers of known deafness genes. In humans, a dominant modifier has been reported to suppress deafness in people homozygous for the DFNB26 recessive deafness locus (Riazuddin *et al.*, 2000). Two modifiers have been proposed to influence the hearing loss observed in carriers of the A1555G mitochondrial mutation associated with progressive hearing loss, a locus on chromosome 8 near D8S277 and a locus on chromosome 6 near the *TFB1M* gene (Bykhovskaya *et al.*, 2000; 2004). In the mouse, three modifiers of deafness genes have been identified. The first is *moth1*, which has a major effect upon the hearing of tubby mutant homozygotes (*tub*/*tub*), and has been identified as a variant of the *Mtap1* gene. A second is *mdfw*, modifier of deaf waddler, which largely determines whether heterozygous carriers of a deaf waddler mutation (+/$Atp2b2^{dfw2J}$) show progressive hearing loss or not. This modifier is most likely to be a synonymous single nucleotide polymorphism that leads to in-frame skipping of exon 7 of the *Cdh23* gene (Noben-Trauth *et al.*, 2003). This variant of *Cdh23* is also known as *Ahl* because it was originally mapped as a locus involved in age-related hearing loss in a number of common mouse strains. The third mouse modifier is a mitochondrial mutation that affects hearing impairment in mice homozygous for the susceptible allele of *Ahl* (*ahl*/*ahl*) and has been identified as a single nucleotide insertion in the *tRNA-Arg* gene.

A model for human hearing impairment?

Age-related hearing loss is an extremely common human disease, so what have we learnt so far that might suggest the causes? Does the common disease/common variant model look the most likely model?

If we look firstly at childhood deafness, despite the large number of genes that have been identified, most families with a deaf child still have no molecular diagnosis. The gene most commonly involved in severe or profound childhood deafness is *GJB2*, encoding the connexin 26 molecule, a component of gap junctions thought to be important for homeostatic control within the cochlear duct. Mutations in this gene account for between 25 and 50% of cases of known recessive congenital severe or profound deafness, and about 10% of sporadic cases. There are still a large number of cases with severe or less severe hearing impairment for which no gene is known. Many of the genes known to be involved in deafness affect only a handful of families. These observations suggest that childhood deafness is probably due to a large number of single gene mutations in different genes, rather than a small set of common polymorphisms, together with some environmental factors such as neonatal problems or infection.

In age-related hearing loss, we know of some single gene mutations that lead to progressive hearing loss with adult onset, mostly dominantly inherited. We know that heritability is relatively high for hearing loss, especially amongst those with onset in middle age. We also know that childhood deafness is extremely heterogeneous, with many different genes involved. All this argues in favour of a similar model for age-related hearing loss as the simplest explanation: many different single genes, rather than a limited number of common gene variants. Extreme heterogeneity would make straightforward case/control association studies difficult. The National Study of Hearing carried out in the UK twenty years ago failed to find association of any biomarkers studied with hearing impairment, which is consistent with extensive heterogeneity. The genes involved with age-related hearing loss could overlap with those involved in childhood deafness, because it is easy to imagine that a milder mutation could lead to later-onset auditory pathology. However, we also know that there are complex interactions that influence

auditory function. Firstly, some genes act as modifiers of other genes in determining hearing ability. Secondly, genes interact with environmental factors making some people more or less sensitive to auditory insults. Thirdly, hearing ability may be determined by the summed action of a number of gene variants acting as quantitative trait loci (QTLs). There is evidence from the mouse that the progressive hearing loss observed in many inbred strains of mouse is due to a combination of several QTLs, including the *ahl* variant of *Cdh23* on chromosome 10 mentioned earlier, the *Ahl2* locus on chromosome 5, and *Ahl3* on chromosome 17 (Noben-Trauth *et al.*, 2003), as well as further QTLs not yet published.

Whichever model explains age-related hearing loss, it is likely that the genes involved will include those that have a specific influence on hearing and not just those involved in systemic aging processes that affect other senses and neural function (see **C**hapter 8). Looking directly for linked or associated variants in candidate genes suggested by study of early-onset deafness in humans and mouse models may be a useful approach for age-related hearing loss. Choice of the population will be important, and initially individuals at the extremes of the distribution (in terms of age of onset or rate of progression) might be most fruitful (Wright *et al.*, 2003). Once the major genes involved in hearing loss in the human population have been identified, work on understanding their role and developing treatments can be more focussed.

Acknowledgements

Supported by the Wellcome Trust, MRC, Defeating Deafness and EC.

REFERENCES

Anagnostopoulos, A.V. (2002). A compendium of mouse knockouts with inner ear defects. *Trends Genet.*, **18**, 499.

Bitner-Glindzicz, M. (2002). Hereditary deafness and phenotyping in humans. *Br Med Bull.*, **63**, 73–94.

Browning, G.G. and Gatehouse, S. (1992). The prevalence of middle ear disease in the adult British population. *Clin Otolaryngol.*, **17**, 317–21.

Bykhovskaya, Y. *et al.* (2000). Candidate locus for a nuclear modifier gene for maternally inherited deafness. *Am J Hum Genet.*, **66**, 1905–10.

Bykhovskaya, Y. *et al.* (2004). Human mitochondrial transcription factor B1 as a modifier gene for hearing loss associated with the mitochondrial A1555G mutation. *Mol Genet Metab.*, **82**, 27–32.

Christensen, K. *et al.* (2001). Genetic and environmental influences on self-reported reduced hearing in the old and oldest old. *J Am Geriatr Soc.*, **49**, 1512–17.

Coling, D.E. *et al.* (2003). Effect of *SOD1* overexpression on age- and noise-related hearing loss. *Free Radic Biol Med.*, **34**, 873–80.

Davis, A.C. (1989). The prevalence of hearing impairment and reported hearing disability among adults in Great Britain. *Int J Epidemiol.*, **18**, 911–17.

Davis, A.C. and Moorjani, P. (2002). The epidemiology of hearing and balance disorders. In Luxon, L.M. (ed.), *Textbook of audiological medicine: clinical aspects of hearing and balance*, pp. 89–99. London: Martin Dunitz.

Davis, R.R. *et al.* (2002). Susceptibility to the ototoxic properties of toluene is species specific. *Hear Res.*, **166**, 24–32.

DeStefano, A.L. *et al.* (2003). Genomewide linkage analysis to presbycusis in the Framingham Heart Study. *Arch Otolaryngol Head Neck Surg.*, **129**, 285–9.

Fechter, L.D. (2004). Promotion of noise-induced hearing loss by chemical contaminants. *J Toxicol Environ Health A.*, **67**, 727–40.

Fortnum, H.M. *et al.* (2001). Prevalence of permanent childhood hearing impairment in the United Kingdom and implications for universal neonatal hearing screening: questionnaire based ascertainment study. *BMJ*, **323**, 536–40.

Fortnum, H.M. *et al.* (2002). Epidemiology of the UK population of hearing-impaired children, including characteristics of those with and without cochlear implants–audiology, aetiology, comorbidity and affluence. *Int J Audiol.*, **41**, 170–9.

Fransen, E. *et al.* (2003). Age-related hearing impairment (ARHI): environmental risk factors and genetic prospects. *Exp Gerontol.*, **38**, 353–9.

Friedman, T. B. and Griffith, A. J. (2003). Human nonsyndromic sensorineural deafness. *Annu Rev Genomics Hum Genet.*, **4**, 341–402.

Gates, G. A. *et al.* (1999). Genetic associations in age-related hearing thresholds. *Arch Otolaryngol Head Neck Surg.*, **125**, 654–9.

Hakuba, N. *et al.* (2000). Exacerbation of noise-induced hearing loss in mice lacking the glutamate transporter GLAST. *J Neurosci.*, **20**, 8750–3.

Holme, R. H. and Steel, K. P. (2004). Progressive hearing loss and increased susceptibility to noise-induced hearing loss in mice carrying a *Cdh23* but not a *Myo7a* mutation. *J Assoc Res Otolaryngol.*, **5**, 66–79.

Huygen, P. L. M. and Verhagen, W. I. M. (1994). Peripheral vestibular and vestibulo-cochlear dysfunction in hereditary disorders. *J. Vestibular Research*, **4**, 81–104.

Johansen, I. R. *et al.* (2004). Longitudinal study of hearing impairment in children. *Int J Pediatr Otorhinolaryngol.*, **68**, 1157–65.

Karlsson, K. K. *et al.* (1997). Description and primary results from an audiometric study of male twins. *Ear Hear.*, **18**, 114–20.

Kozel, P. J. *et al.* (2002). Deficiency in plasma membrane calcium ATPase isoform 2 increases susceptibility to noise-induced hearing loss in mice. *Hear Res.*, **164**, 231–9.

Maison, S. F. *et al.* (2002). Efferent protection from acoustic injury is mediated via alpha9 nicotinic acetylcholine receptors on outer hair cells. *J Neurosci.*, **22**, 10838–46.

Morzaria, S. *et al.* (2004). Systematic review of the etiology of bilateral sensorineural hearing loss in children. *Int J Pediatr Otorhinolaryngol.*, **68**, 1193–8.

Nishi, M. *et al.* (1997). Unrestrained nociceptive response and disregulation of hearing ability in mice lacking the nociceptin/orphaninFQ receptor. *EMBO J.*, **16**, 1858–64.

Noben-Trauth, K. *et al.* (2003). Association of cadherin 23 with polygenic inheritance and genetic modification of sensorineural hearing loss. *Nat Genet.*, **35**, 21–3.

Ohlemiller, K. K. *et al.* (1999). Targeted deletion of the cytosolic Cu/Zn-superoxide dismutase gene (*Sod1*) increases susceptibility to noise-induced hearing loss. *Audiol Neurootol.*, **4**, 237–46.

Ohlemiller, K. K. *et al.* (2000). Targeted mutation of the gene for cellular glutathione peroxidase (*Gpx1*) increases noise-induced hearing loss in mice. *J Assoc Res Otolaryngol.*, **1**, 243–54.

Peng, B. G. *et al.* (2004). Acid-sensing ion channel 2 contributes a major component to acid-evoked excitatory responses in spiral ganglion neurons and plays a role in noise susceptibility of mice. *J Neurosci.*, **24**, 10167–75.

Petit, C. *et al.* (2001). Molecular genetics of hearing loss. *Annu Rev Genet.*, **35**, 589–646.

Riazuddin, S. *et al.* (2000). Dominant modifier DFNM1 suppresses recessive deafness DFNB26. *Nat Genet.*, **26**, 431–4.

Rosenhall, U. (2002). Ageing in the auditory and vestibular systems. In Luxon, L.M. (ed.), *Textbook of audiological medicine: clinical aspects of hearing and balance*, pp. 14–23. London: Martin Dunitz.

Rosowski, J. J. *et al.* (2003). The aging of the middle ear in *129S6/SvEvTac* and *CBA/CaJ* mice: measurements of umbo velocity, hearing function, and the incidence of pathology. *J Assoc Res Otolaryngol.*, **4**, 371–83.

Schick, B. *et al.* (2004). Increased noise sensitivity and altered inner ear MENA distribution in *VASP-/-* mice. *Cell Tissue Res.*, **318**, 493–502.

Schuknecht, H. F. and Gacek, M. R. (1993). Cochlear pathology in presbycusis. *Ann Otol Rhinol Laryngol.*, **102**, 1–16.

Steel, K. P. (2000). Science, medicine, and the future – new interventions in hearing impairment. *BMJ*, **320**, 622–5.

Sugahara, K. *et al.* (2003). Heat shock transcription factor HSF1 is required for survival of sensory hair cells against acoustic overexposure. *Hear Res.*, **182**, 88–96.

Van Camp, G. *et al.* (1997). Nonsyndromic hearing impairment: unparalleled heterogeneity. *Am J Hum Genet.*, **60**, 758–64.

Van Camp, G. and Smith, R. (2005). The Hereditary Hearing Loss Homepage. www.webhost.ua.ac.be/hhh/

Vázquez, A. E. *et al.* (2004). Evaluating cochlear function and the effects of noise exposure in the *B6.CAST + Ahl* mouse with distortion product otoacoustic emissions. *Hear Res.*, **194**, 87–96.

Wang, J. *et al.* (2003). A peptide inhibitor of c-Jun N-terminal kinase protects against both aminoglycoside and acoustic trauma-induced auditory hair cell death and hearing loss. *J Neurosci.*, **23**, 8596–607.

Wright, A. *et al.* (2003). A polygenic basis for late-onset disease. *Trends Genet.*, **19**, 97–106.

Yoshida, N. and Liberman, M. C. (2000). Sound conditioning reduces noise-induced permanent threshold shift in mice. *Hear Res.*, **148**, 213–19.

Yoshida, N. *et al.* (1999). Heat stress and protection from permanent acoustic injury in mice. *J Neurosci.*, **19**, 10116–24.

Yoshida, N. *et al.* (2000). Acoustic injury in mice: 129/SvEv is exceptionally resistant to noise-induced hearing loss. *Hear Res.*, **141**, 97–106.

Zhai, S. Q. *et al.* (2004). Basic fibroblast growth factor protects auditory neurons and hair cells from glutamate neurotoxicity and noise exposure. *Acta Otolaryngol.*, **124**, 124–9.

34

Pharmacogenomics: clinical applications

Gillian Smith, Mark Chamberlain and C. Roland Wolf

Introduction

There is now a large body of evidence indicating that genetic factors can influence treatment response in a variety of human diseases. Pharmacogenetics, defined as the study of the genetic basis for individuality in response to drugs (Vogel, 1959) or use of genetic analysis to predict drug response, efficacy and toxicity (Roses, 2004) has seen an almost exponential rise in peer-reviewed publications in the last 10 years (Figure 34.1). The closely related field of pharmacogenomics can be defined as the use of genetic approaches in drug discovery and use and encompasses the study of all genes which may influence drug response although, as illustrated in Figure 34.1, the terms pharmacogenetics and pharmacogenomics are often used interchangeably.

Inter-individual differences in response to commonly prescribed drugs are increasingly recognized. Serious adverse drug reactions (ADRs) have been estimated to occur in more than 2 million patients/year in the USA and result in more than 100 000 fatalities; meta-analysis of a number of US-based prospective studies suggest that ADRs are the fourth leading cause of death, after heart disease, cancer and stroke (Lazarou *et al.*, 1998). A recent UK-based prospective study reported that up to 6.5% of all hospital admissions in the UK are related to adverse drug reactions, with a projected annual cost to the National Health Service of £466 million (Pirmohamed *et al.*, 2004).

Individuality in drug response can be both genetically and environmentally determined. We know, for example, that the genes that determine the metabolism, disposition and often the toxicity of drugs are regulated by a large number of endogenous and exogenous factors. Therefore exposure to such agents, in the diet for example, or by interactions with other drugs prescribed concomitantly, can have profound effects on circulating drug levels or the levels of toxic metabolites. In addition, we know that genetic polymorphisms can have a profound effect on drug outcome as well as on drug pharmacokinetics. Where the balance lies between environmental and genetic variability, however, remains poorly understood. Although there have now been numerous pharmacogenetic studies, the value of many of these has been reduced because of difficulties in defining response or non-response, and many studies have lacked significant statistical power. As a consequence, there remain many contradictions on this theme within the literature. However, recent advances in biotechnology, bioinformatics and genetic information (Table 34.1) continue to increase our ability to obtain detailed phenotypic information for individual patients, which can be used to predict response or side effect profiles following drug treatment and creates exciting new possibilities to carry out pharmacogenetic research in a more meaningful way.

Although pharmacogenetics was first described in the late 1950s (Vogel, 1959) and DNA-based tests to predict drug response have been

Table 34.1. Databases for pharmacogenetic variation

ALFRED (allele frequency database)- a resource of gene frequency data on human populations supported by the US National Science Foundation: www.alfred.med.yale.edu/alfred/
EMBL-EBI – 'SVD' sequence variation database project: www2.ebi.ac.uk/mutations
Ensembl – joint project between Sanger institute and EBI: www.ensembl.org/
HGMD – human gene mutation database: www.archive.uwcm.ac.uk/uwcm/mg/hgmd0.html?
HGVS – human genome variation society: www.genomic.unimelb.edu.au/mdi/dblist/dblist.html
Human genome variation database (HGVbase): http://hgvbase.cgb.ki.se/
Human Cytochrome P450 (*CYP*) Allele Nomenclature Committee: www.imm.ki.se/CYPalleles/
National Centre for Biotechnology Information (NCBI) dbSNP: www.ncbi.nlm.nih.gov/SNP/
PharmGKB – an integrated resource about how variation in human genes leads to variation in our response to drugs: http://www.pharmgkb.org/
SNP Consortium Ltd – SNPs for biomedical research and dbs for mutations in human genes: www.umd.necker.fr/

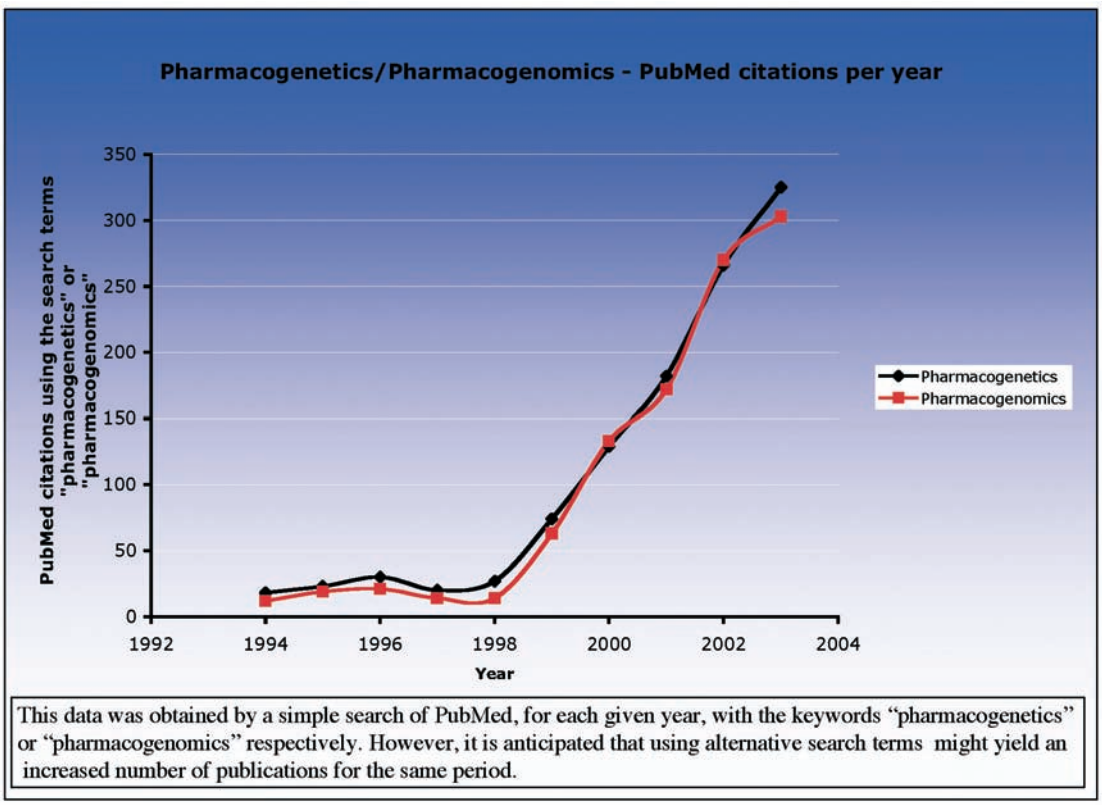

Figure 34.1 Pharmacogenetics/Pharmacogenomics – PubMed citations per year.

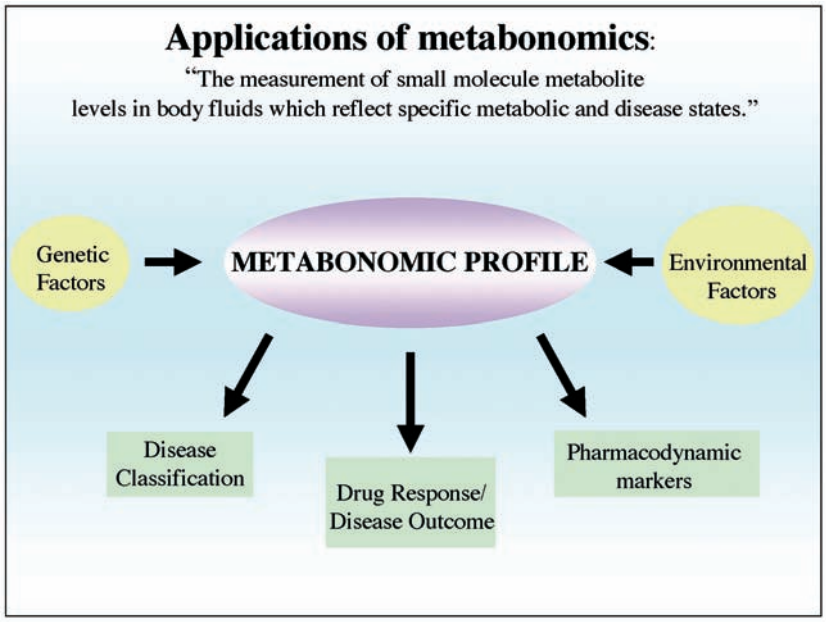

Figure 34.2 Applications of metabonomics: ''The measurement of small molecule metabolite levels in body fluids which reflect specific metabolic and disease states.''

available for the last 15 years, it is only relatively recently that the technology and information on pathways of drug metabolism has become sufficiently advanced to allow multiple candidate genes to be studied simultaneously or to carry out genome-wide screens to identify genes that influence drug response. While genotype can be determined unequivocally, one of the most significant current limitations in the development of this research area is finding better ways of rapidly defining phenotype in large populations. This is particularly the case, as discussed below, in the treatment of psychiatric illness, where definitions of response and non-response are subjective and are based on detailed psychiatric interviews. Pharmacogenetic testing by DNA-based analysis alone in most cases is unlikely to accurately predict patient phenotype and therefore response to drug treatment. In the majority of cases, assessment of patient phenotype is complex and does not arise from genetically determined variation in a single gene.

Recent developments in comprehensive metabolite profiling or ''metabonomics'' may lead to significant improvements in phenotype prediction (Nicholson and Wilson, 2003). Metabonomics provides a ''real time'' assessment of individuality in response to all endogenous and exogenous stimuli including drug treatments and, unlike pharmacogenetics, allows for inter-individual differences in, for example, liver and renal function, disease activity, concomitant treatment and drug/drug interactions. Metabonomic profiling, using sophisticated and highly sensitive liquid chromatography and mass spectrometry techniques, permits the simultaneous analysis not only of urinary drug and metabolite profiles but also of metabolites derived from endogenous chemicals and their breakdown products (Figure 34.2).

Genes that influence drug response

Many of the early studies on pharmacogenetics have focussed on polymorphisms in drug metabolising enzymes, most notably within Phase I enzymes such as the cytochrome P450s. Numerous alleles in the major drug metabolising P450s have now been reported (Table 34.1, summarized at the website, www.imm.ki.se/CYPalleles/). The majority of the research effort to date has focussed on genes within the *CYP2C*, *CYP2D* and *CYP3A* families, which together account for the majority of Phase I drug metabolism in man and which show marked inter-individual differences in expression and activity (Shimada *et al.*, 1994; Forrester *et al.*, 1992). A recent analysis of the extent of genetically-determined variation in 11 ''drug metabolizing'' P450 genes in families *CYP1*, *CYP2* and *CYP3* in an ethnically diverse population identified a total of 388 polymorphisms, 74 (19%) of which represented coding sequence changes (Solus *et al.*, 2004), and it has been suggested that individuality in P450 activity may account for up to 20% of all variability in response to drug treatment (Ingelman-Sundberg, 2004).

Individuality in *CYP2C9* expression and activity has predominantly been correlated with inheritance of the *CYP2C9*2* and *CYP2C9*3* alleles, each of which contains a single amino acid substitution compared to the consensus sequence, *CYP2C9*1* (Stubbins *et al.*, 1996). Both alleles have been associated with reduced enzyme activity and, although the effects are relatively modest and substrate specific, there is now consistent evidence that *CYP2C9* genotype, particularly inheritance of the *CYP2C9*3* allele, influences sensitivity to drugs including warfarin (reviewed by Daly and Aithal, (2003)) and ibuprofen (Garcia-Martin *et al.*, 2004). However, as the *CYP2C9*3* allele is relatively rare, present in approximately 8% of the Caucasian population, it is unlikely that inheritance of this or other more recently described rarer *CYP2C9* alleles (Goldstein, 2002; Blaisdell *et al.*, 2004) will be major determinants of individuality in response to warfarin therapy.

In contrast, *CYP2D6* has a number of clearly defined phenotypes resulting from the inheritance of either ''poor metaboliser (PM)'' or loss of function alleles (Gonzalez *et al.*, 1988; Gough *et al.*, 1990) or multiple copies of the *CYP2D6* gene, resulting in an ''ultrarapid metaboliser (UM)'' phenotype (Johansson *et al.*, 1993). PM individuals are unable to metabolise drugs which are *CYP2D6* substrates while, in contrast, ultrarapid metabolisers require significantly increased drug doses to achieve the desired therapeutic effect (Bertilsson *et al.*, 1993). More than 90 *CYP2D6* alleles have now been described (www.imm.ki.se/CYPalleles/cyp2d6.htm), the majority of which are rare in White populations and do not have a clearly defined phenotype. However, approximately 7% of Caucasian individuals inherit the CYP2D6 PM phenotype and 2% the UM phenotype (Sachse *et al.*, 1997). This has significant implications for individuality in drug metabolism and disposition, as *CYP2D6* has been estimated to metabolize more than 25% of all commonly prescribed drugs and the majority of drugs used in the treatment of psychiatric disease.

CYP3A4 is the most abundant P450 in human liver and small intestine and is therefore the major drug metabolizing P450, responsible for the metabolism of more than 50% of drugs (Shimada *et al.*, 1994). Inter-individual differences in CYP3A4 expression and activity have been extensively documented (Schuetz *et al.*, 1994; Koch *et al.*, 2002) although, unlike *CYP2D6*, there is no clear genetic basis for individuality in *CYP3A4* activity. To date, more than 40 *CYP3A4* alleles have been described (www.imm.ki.se/CYPalleles/cyp3a4.htm), but the majority of these are rare, or have not been associated with an altered phenotype. Recently, a novel *CYP3A4* promoter polymorphism which was reported to influence gene transcription was described, more than 11 kb upstream of the transcription start site (Matsumura *et al.*, 2004). However, although this study illustrates the benefits of including extended promoter sequences in routine polymorphism screens, this novel allele was found in only 3% of the French

White population suggesting that it is unlikely to be a major determinant of individuality in *CYP3A4* expression.

Individuality in the expression and regulation of *CYP3A* genes is further complicated by a number of factors. *CYP3A4* is a highly inducible gene, where protein expression and therefore catalytic activity is influenced by both endogenous and exogenous chemicals, including concomitantly prescribed drugs (Raucy, 2003; Guengerich, 1999). Many drugs metabolised by *CYP3A4* are also substrates for the drug efflux pump P-glycoprotein, encoded by the *MDR1* gene (Engman *et al.*, 2001; Patel and Mitra, 2001). As *MDR1* is also highly polymorphic (Woodahl and Ho, 2004), it is often not clear whether inter-individual differences in the metabolism of drugs which are substrates for both *CYP3A4* or *MDR1* result from genetically determined differences in the activity of one or both genes, or from additional polymorphisms in transcription factors e.g. the pregnane X receptor (PXR) (Zhang *et al.*, 2001; Hustert *et al.*, 2001) which regulate the expression of both *CYP3A4* and *MDR1*.

CYP3A4 has very similar substrate specificity to *CYP3A5*, a highly homologous gene which lies in close proximity to *CYP3A4* on chromosome 7q21 (Spurr *et al.*, 1989). *CYP3A5* expression is highly variable as a consequence of a polymorphic splice site variant in intron 3 which creates a premature stop codon (Kuehl *et al.*, 2001). *CYP3A5* is expressed in approximately 30% of Whites, and may therefore make a significant contribution to *CYP3A*-catalyzed drug metabolism in these individuals, particularly in extra-hepatic tissues where *CYP3A5* is preferentially expressed.

Other drug metabolizing enzymes for which pharmacogenetic polymorphisms have been correlated with inter-individual differences in expression and activity include the Phase II enzymes N-acetyl transferase 2 (*NAT2*) where ''slow acetylator'' alleles with reduced ability to metabolize drugs such as the anti-tubercular drug isoniazid and the sulphonamide antibiotics are inherited by up to 40% of the White population (Blum *et al.*, 1991). Thiopurine methyl transferase (TPMT) shows genetically determined differences in expression which can have profound consequences for the metabolism of thioguanine drugs used in the treatment of childhood leukaemias (Evans, 2004).

Although many polymorphic drug metabolising enzymes have now been described, the majority of the allelic variants described to date are either relatively rare or do not have a clearly defined phenotype and are therefore unlikely to be major determinants of individuality in drug response. Recent research efforts are therefore becoming increasingly focussed on the analysis of genetic variation in, for example, receptors such as the *Ah* receptor, the pregnane X receptor (*PXR*) and the constitutive androstane receptor (*CAR*) which regulate P450 expression, in drug transporters which regulate active drug transport both into and out of the cell and in genes in downstream signalling pathways which are activated by drug treatment. As the majority of commonly prescribed drugs are, however, rarely substrates for a single enzyme, there are still many complexities associated with understanding and predicting inter-individual differences in drug response.

Clinical applications of pharmacogenetics

The use of pharmacogenetic analysis to predict drug response has applications in all areas of clinical medicine. There are, however, particular examples of diseases such as psychiatric illness and cancer, where individuality in response to drug treatment can be both unpredictable and treatment-limiting.

As many as 40% of patients fail to respond to initial therapy for psychiatric illnesses such as depression (Sackeim, 2001). Although much of this variation in response is likely to reflect inherent difficulties in making unequivocal diagnoses and assessing treatment response, genetically determined individuality in drug targets is known to significantly influence treatment response. For example, the majority of drugs used

Table 34.2. Summary of preliminary average dose recommendations for *CYP2D6* dependent antidepressant drugs (single dose and maintenance treatment)

Drug	M/S	Usual dose(mg)	UM	EM	IM	PM
Amitriptyline	M	150 (50–150)*		120%	(90%)	50%
	S	50 (50–150)*		120%	80%	70%
Clomipramine	M	150 (100–200)*		120%	(90%)	60%
	S	50 (100–200)*		120%	(90%)	60%
Desipramine	M	150 (10–100)*		130%	30%	30%
	S	50 (10–100)*	260%	130%	80%	20%
Fluvoxamine	M	100 (100)*		110%	(100%)	90%
	S	50 (50)*		120%	(90%)	60%
Imipramine	M	150 (25–100)*		130%	(80%)	30%
	S	50 (25–100)*		110%	(100%)	60%
Nortriptyline	M	150 (25–150)*		120%	90%	50%
	S	50 (25–150)*	230%	140%	70%	50%
Paroxetine	M	20 (30)*		110%	(90%)	70%
	S	20 (30)*		130%	(80%)	20%

M/S recommendations for multiple-dosing (M, maintenance treatment) or for beginning of treatment (S, single dose). Usual dose is dose usually given to patients. ( )* = range of dose given in pharmacogenetic studies.
( ) = estimations for IMs bsased on analogy. UM = ultra rapid metabolizer; EM = extensive metabolizer; IM = intermediate metabolizer; PM = poor metabolizer.
Adapted from: Kirchheiner *et al.*, *Acta Psychiatr Scand*, 2001, **104**, 173–92.

in the treatment of psychiatric illness are small lipophilic molecules which are substrates for the polymorphic enzyme *CYP2D6* (Kirchheiner *et al.*, 2004). *CYP2D6* genotype can therefore influence the bioavailability and efficacy of drugs such as the selective serotonin reuptake inhibitors (SSRI) paroxetine (Seroxat) and fluoxetine (Prozac). As a consequence, prescribing guidelines have recently been proposed, where ''standard'' dosing schedules are modified according to *CYP2D6* genotype (Kirchheiner *et al.*, 2001; 2004) (Table 34.2). In addition to polymorphisms in drug metabolising enzymes, there is increasing evidence that genetically determined differences in drug receptors and transporters (e.g. the post-synaptic serotonin receptors and serotonin re-uptake transporter) and in downstream signalling pathways can influence treatment response (Lerer and Macciardi, 2002).

Cancer chemotherapy is another area in which pharmacogenetic polymorphisms can have a very significant influence on treatment response. As the majority of anti-cancer drugs are by definition highly cytotoxic agents with narrow therapeutic indices, even subtle changes in drug metabolizing enzyme activity can have profound consequences. There is increasing evidence that pharmacogenetic polymorphisms can influence treatment response. For example, it has recently been demonstrated that glutathione S-transferase P1 (*GSTP1*) genotype influences survival following the treatment of colorectal cancer with a combined 5-FU and oxaliplatin regimen (Stoehlmacher *et al.*, 2002; Stoehlmacher *et al.*, 2004), while polymorphisms in the UDP-glucuronosyl transferase 1A1 (*UGT1A1*) gene have been shown to influence response to irinotecan in colorectal cancer patients (Marsh and McLeod, 2004).

Response to chemotherapy is also influenced by individuality in tumor gene expression, where variability in the target cells quite clearly

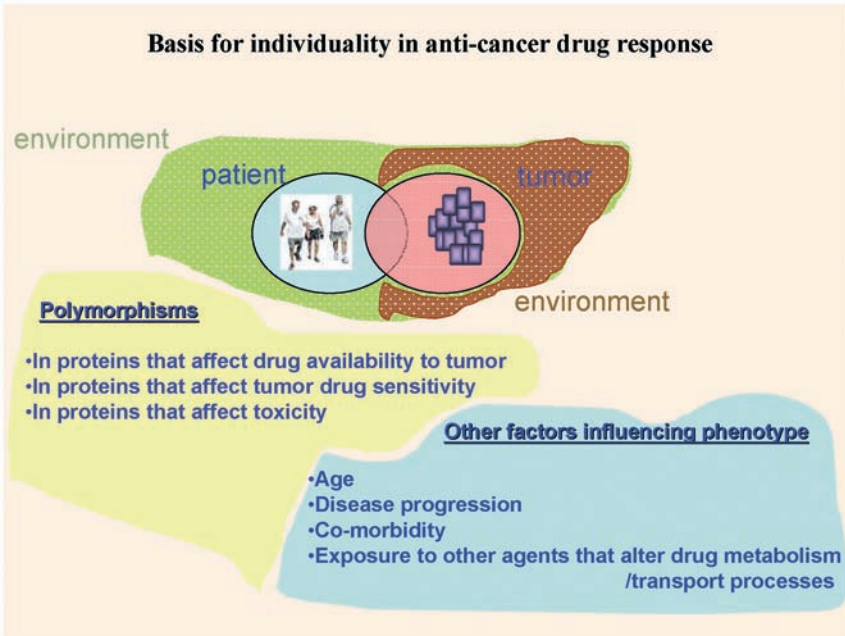

Figure 34.3 Basis for individuality in anti-cancer drug response.

influences whether or not an individual responds to a particular cytotoxic agent (Figure 34.3). For example, Iressa (Gefitinib) is a relatively new drug which has recently been approved in the USA for the treatment of non-small-cell lung cancer (Wakeling *et al.*, 2002). Iressa inhibits the tyrosine kinase activity of the epidermal growth factor receptor (*EGFR*), a gene which is frequently over-expressed or activated by mutations in a variety of human cancers, but is effective in only approximately 10% of patients (Fukuoka *et al.*, 2003). Analysis of tumors from ''responders'' and ''non-responders'' following Iressa treatment revealed the presence of somatic *EGFR* mutations in the response group only, suggesting that analysis of EGFR mutation status prior to treatment may be predictive of response (Paez *et al.*, 2004). Interestingly, it has recently been reported that EGFR mutations and therefore response to Iressa is greater in non-smokers (Pao *et al.*, 2004), highlighting the critical interaction between genetic background and environmental exposure in determining drug response. Similar to Iressa, Herceptin (Trastuzumab) is a humanized monoclonal antibody raised against the *ERBB2* tyrosine kinase receptor which is over-expressed in up to 30% of breast tumors (Bell *et al.*, 2004). Although Herceptin was initially licensed for the treatment of all breast cancer, it is now appreciated that treatment response is increased in patients over-expressing *ERBB2*. We have recently also demonstrated marked inter-individual differences in *APC, K-ras* and *p53* tumor genotype in patients with colorectal cancer (Smith *et al.*, 2002), suggesting that both patient and tumour gene expression may be a significant determinant of treatment response.

Ethnic differences in allele frequencies

Marked ethnic differences in allele frequencies for the majority of the major drug metabolizing enzymes have been documented (Kim *et al.*, 2004). Ethnic diversity in allele frequencies is

Table 34.3. *CYP2D6*: Ethnic diversity

Allele frequencies *CYP2D6*	Africans	African Americans	Asians	Whites
CYP2D6 1/* 2*	67%	62%	55%	73%
CYP2D6 3*	0%	0.3%	n.d.	1.2%
CYP2D6 4*	3.9%	7.8%	0.2%	17%
CYP2D6 5*	4.7%	6.2%	4.5%	3.4%
CYP2D6 6*	0%	n.d.	n.d.	1.0%
CYP2D6 7*	0%	n.d.	n.d.	0.1%
CYP2D6 8*	0%	n.d.	n.d.	0.1%
CYP2D6 9*	0%	n.d.	n.d.	1.9%
CYP2D6 10*	6.7%	7.5%	38%	2.9%
CYP2D6 29*	20%	n.d.	n.d.	0.1%
CYP2D6 35*	n.d.	n.d.	n.d.	6.7%
CYP2D6 1 × 2*	4.2%	n.d.	0.5%	0.8%
CYP2D6 2 × 2*	6.7%	1.9%	0.5%	1.7%
CYP2D6 4 × 2*	0.9%	n.d.	n.d.	0.5%
CYP2D62 × N	n.d.	n.d.	1%	1.9%

particularly well characterized for *CYP2D6*, as illustrated in Table 34.3. While the ''poor metaboliser'' *CYP2D6* phenotype is found in approximately 7% of White populations, it is rare (frequency of 1–2%) in Asian and Oriental populations (Lou *et al.*, 1987). In contrast, certain *CYP2D6* alleles associated with reduced enzyme activity (e.g. *CYP2D6* 10*) are much more common or are exclusively found in Oriental populations (Johansson *et al.*, 1994). A particularly interesting example of ethnic diversity in drug metabolism is the *CYP2D6* gene duplication, which is relatively common (allele frequency ~30%) in African and Arabic populations, is found at a frequency of 10% in Spaniards and in Italian and Turkish populations (Agundez *et al.*, 1994), but only at a frequency of 1–2% in Northern Europeans and is essentially absent in Asians (Ingelman-Sundberg *et al.*, 1999). It has been suggested that as *CYP2D6*, unlike other *P450s*, is not a highly inducible gene, the gene duplication may have arisen as an evolutionary adaptive response to environmental challenge on exposure to toxic alkaloids in the African diet (Aklillu *et al.*, 2002).

Although much of the work to date on ethnic differences in allele frequencies has focussed on drug-metabolizing enzymes, there is increasing evidence that there are many other examples of genes where ethnic differences in allele (or tumor mutation frequencies) can have significant implications for drug response. For example, as discussed above, the lung cancer drug Iressa (Gefitinib) is only active in chemotherapy-resistant patients with mutations in the *EGFR* gene. While some 10% of Caucasian lung cancer patients respond to Iressa, up to 28% of Japanese patients have been reported to have *EGFR* mutations and to respond to similar treatment schedules (Fukuoka *et al.*, 2003).

As the majority of pharmacogenetic studies to date have been carried out in White populations, there is an urgent need to expand these studies to encompass all ethnic groups. The development of SNP and haplotype maps will facilitate these studies; however, they will only be feasible in populations where the drugs are actually prescribed. Such studies, in African countries for example, may be extremely difficult. The increasing heterogeneity of the human population implies that the implications of genetic variability as they relate to drug response must be understood across the entire population.

Current applications of pharmacogenetics

Our knowledge of pharmacogenetics can be applied in two main areas – (a) in improving the use of existing commonly prescribed drugs and (b) in the development and use of new drugs by the pharmaceutical and biotechnology industries.

a) Current drugs

The ability to individualize the use of currently used drugs will not only permit the identification of individuals vulnerable to drug side effects but, perhaps more importantly, will allow us to identify those patients who will or will not respond to a particular drug treatment. As discussed above, there are certain research areas where this is of vital importance, e.g. in the treatment of cancer, but also for other diseases such as the control of serum cholesterol with statins or the use of non-steroidal anti-inflammatory drugs, where a significant proportion of the population do not respond to these agents. For certain drugs there are already prescribing guidelines; for example, in the treatment of childhood leukaemia with 6-mercaptopurine (Armstrong *et al.*, 2004) and in the use of Glivec (Vastag, 2004) and Herceptin (Bell *et al.*, 2004). There is, however, still an urgent need for new studies to identify genes which influence drug metabolism and efficacy, and to correlate genetically determined variation in these genes with treatment response. It is, of course, possible that such pharmacogenetic studies could be incorporated within other existing large population-based studies such as deCODE (www.decode.com) and the UK Biobank (www.ukbiobank.ac.uk). However, unless the requisite protocols and infrastructure for pharmacogenetic analysis are embedded in the conceptual and design phases of these studies it will be impossible to collect this information retrospectively.

Although pharmacogenetic approaches to increase our understanding of individuality in drug response have enormous implications for routine medical practice, studies on currently used drugs are unlikely to be sponsored by the pharmaceutical industry and it therefore increasingly becomes the responsibility of central Government to support this type of research at an international level. One way to carry out these studies effectively is to use a population-based approach where pharmacogenetic factors that determine patient response can be correlated with genetic background and detailed computer-based information on drug prescribing and outcome in individual patients.

b) The development of new drugs

The pharmaceutical industry is acutely aware of the need to consider the implications of pharmacogenetics on the discovery and use of new drugs. It has important implications for the industry as, in the first instance, many new drug targets are identified as susceptibility genes and are therefore by definition polymorphic in the human population. A number of major pharmaceutical companies now routinely screen new drug targets to establish the extent to which genetically determined variation is of significance to new drug development programmes. If a drug target is polymorphic, this information may be used to identify individuals who will respond differently to drug treatment. Although this information is not currently publicly available, new FDA (Food and Drug Administration) guidelines in the USA will require the inclusion of pharmacogenetic information as part of future regulatory submissions for all new drugs (Lesko and Woodcock, 2002).

The most appropriate time to establish whether a polymorphism in a new drug target is an important determinant of therapeutic efficacy is during Phase 3 clinical trials, where blood samples are routinely collected from all trial participants and where the incorporation of additional genotype analysis will therefore be relatively routine and will not have significant cost implications. The ability of pharmaceutical companies to use genetic information to subdivide their patient populations into responders and non-responders may, on the one hand, reduce their market share, as it may not be possible to design new ''blockbuster'' drugs suitable for all patients. However, it is also possible that genetic testing may identify a subpopulation of

responders, which would permit limited drug registration. The concept that future new drugs may be licensed together with an appropriate genetic test is attractive, although it would be necessary to limit prescription to areas/countries where both genetic testing and appropriate qualified interpretation is routinely available.

Future developments

Much of the current uncertainty about the application of pharmacogenetics in routine medical care is driven by the limited knowledge of the medical profession about what pharmacogenetics actually is, and how it may be applied in what is already an overburdened healthcare system. Recent advances in the subclassification of disease and, as a consequence, more sophisticated individualization of patient care will undoubtedly become part of routine clinical practice in the future. In certain disease areas, however, there remains a lack of effective drug treatments and patient care remains relatively primitive. This is particularly the case in the treatment of cancer and a number of Third World parasitic diseases, where highly toxic drug regimens and drugs are used with only a relatively small proportion of patients receiving any benefit. For example, many colorectal cancer patients are treated with 5-fluorouracil, where only 10–20% of patients will show significant responses, and the majority will experience treatment-limiting side effects.

High-throughput genotyping technology and the development of SNP (single nucleotide polymorphism) and haplotype maps of the Human Genome means that, in the foreseeable future, it will be possible to carry out genome-wide pharmacogenetic testing of alleles or SNPs quickly and cheaply with minimal cost implications for healthcare providers. There are no practical reasons why the instigation of routine genotyping tests within, for example, clinical biochemistry departments, should not become as routine as radioimmuno and ELISA-based assays. It is, however, important that we are realistic about the extent to which routine pharmacogenetic testing can truly individualize drug treatment. Current enthusiasm in the scientific and medical literature about the applicability of pharmacogenetics suggests that the incorporation of pharmacogenetic testing into routine clinical practice will lead to an ''all or nothing'' situation where drugs become 100% more effective, with 100% fewer side-effects. This is obviously far from realistic, but it is not yet clear what level of increased predictability would justify routine screening of all patients prior to drug prescription. Almost certainly this is an issue that will have to be independently assessed for individual drugs, taking into account clinical utility, therapeutic index and prescribing costs.

Table 34.4. Clinical implications of pharmocogenetics

- Prediction of responders and non-responders to drug treatment
- Selection of the most appropriate drugs and doses for individual patients
- Avoidance of side-effects
- Increased patient compliance
- Saving of doctors' time
- Reduced drug prescribing costs

There is an increasing need to raise awareness among primary care physicians and hospital-based clinicians of the implications of pharmacogenetic testing (Table 34.4). At present, drug prescribing and use is relatively subjective where, if a particular drug does not work, a doctor will either increase the dose or change to an alternative form of treatment, often without any logical rationale for doing so other than apparent lack of patient response. If, in future, we hope to use pharmacogenetics to individualise drug prescribing, this will not be achieved solely by making better drugs, but must involve an increased understanding of genetic and phenotypic differences between individuals. There is therefore an urgent need to incorporate the concepts of pharmacogenetics into the undergraduate medical curriculum.

Ethical issues

Whether or not pharmacogenetic testing requires the same ethical considerations as other forms of genetic testing has been widely debated (Breckenridge *et al.*, 2004). However, whether or not there is a need for patient counseling very much depends on the specific pharmacogenetic tests being carried out and whether these have associated implications for disease susceptibility. Pharmacogenetic tests therefore need to be developed with guidance for doctors as to whether or not a specific test has any additional implications for the patient and whether genetic counseling would therefore be advised. However, even in the area of human genetics, we have not yet come to grips with the concepts and implications of low levels of relative risk. To date, most of the evidence in favour of pharmacogenetic testing has not identified a single SNP or allele that in its own right would justify genetic counseling.

There are more complex ethical issues relating to routine pharmacogenetic testing. For example, if a pharmacogenetic test is positive and a doctor does not alter the choice or dose of a drug in the appropriate manner, should this be considered negligent on their part? There also remains the possibility that over-prescribing of drugs in vulnerable populations may result in altered aspects of human behaviour for which the doctor may be considered responsible.

In conclusion, if we are to improve the way in which we prescribe and use medicines, more sophisticated phenotypic and genetic tests will become an essential part of routine clinical practice. However, in order to instigate specific pharmacogenetic tests it will be necessary to first clearly demonstrate their value. This will require comprehensive clinical trials, on both currently used drugs and on new drugs in development and there is therefore an onus on the pharmaceutical industry and the drug regulatory agencies to understand how genetic variability may influence drug response. Identification of polymorphic drug targets that result in altered efficacy or side-effect profiles has obvious implications for future drug registration and use. There is now an urgent need to introduce pharmacogenetics to the undergraduate medical curriculum and to simplify associated costs, concepts and prescribing guidelines so that they can be much more accessible to medical practitioners.

REFERENCES

Agundez, J. A., Martinez, C., Ledesma, M. C. *et al.* (1994). Genetic basis for differences in debrisoquine polymorphism between a Spanish and other white populations. *Clin Pharmacol Ther*, **55**, 412–17.

Aklillu, E., Herrlin, K., Gustafsson, L. L., Bertilsson, L. and Ingelman-Sundberg, M. (2002). Evidence for environmental influence on CYP2D6-catalysed debrisoquine hydroxylation as demonstrated by phenotyping and genotyping of Ethiopians living in Ethiopia or in Sweden. *Pharmacogenetics*, **12**, 375–83.

Armstrong, V. W., Shipkova, M., von Ahsen, N. and Oellerich, M. (2004). Analytic aspects of monitoring therapy with thiopurine medications. *Ther Drug Monit*, **26**, 220–6.

Bell, R., Verma, S., Untch, M., Cameron, D. and Smith, I. (2004). Maximizing clinical benefit with trastuzumab. *Semin Oncol*, **31**, 35–44.

Bertilsson, L., Dahl, M. L., Sjoqvist, F. *et al.* (1993). Molecular basis for rational megaprescribing in ultrarapid hydroxylators of debrisoquine. *Lancet*, **341**, 63.

Blaisdell, J., Jorge-Nebert, L. F., Coulter, S. *et al.* (2004). Discovery of new potentially defective alleles of human CYP2C9. *Pharmacogenetics*, **14**, 527–37.

Blum, M., Demierre, A., Grant, D. M., Heim, M. and Meyer, U. A. (1991). Molecular mechanism of slow acetylation of drugs and carcinogens in humans. *Proc Natl Acad Sci USA*, **88**, 5237–41.

Breckenridge, A., Lindpaintner, K., Lipton, P. *et al.* (2004). Pharmacogenetics: ethical problems and solutions. *Nat Rev Genet*, **5**, 676–80.

Daly, A. K. and Aithal, G. P. (2003). Genetic regulation of warfarin metabolism and response. *Semin Vasc Med*, **3**, 231–8.

Engman, H. A., Lennernas, H., Taipalensuu, J. *et al.* (2001). CYP3A4, CYP3A5, and MDR1 in human small and large intestinal cell lines suitable for drug transport studies. *J Pharm Sci*, **90**, 1736–51.

Evans, W. E. (2004). Pharmacogenetics of thiopurine S-methyltransferase and thiopurine therapy. *Ther Drug Monit*, **26**, 186–91.

Forrester, L. M., Henderson, C. J., Glancey, M. J. *et al.* (1992). Relative expression of cytochrome P450 isoenzymes in human liver and association with the metabolism of drugs and xenobiotics. *Biochem J*, **281**, 359–68.

Fukuoka, M., Yano, S., Giaccone, G. *et al.* (2003). Multi-institutional randomized phase II trial of gefitinib for previously treated patients with advanced non-small-cell lung cancer. *J Clin Oncol*, **21**, 2237–46.

Garcia-Martin, E., Martinez, C., Tabares, B., Frias, J. and Agundez, J. A. (2004). Interindividual variability in ibuprofen pharmacokinetics is related to interaction of cytochrome P450 2C8 and 2C9 amino acid polymorphisms. *Clin Pharmacol Ther*, **76**, 119–27.

Goldstein, J. A. (2002). CYP2C9 polymorphisms and CYP2C9*2 genotyping primers. *Br J Clin Pharmacol*, **53**, 409–10.

Gonzalez, F. J., Skoda, R. C., Kimura, S. *et al.* (1988). Characterization of the common genetic defect in humans deficient in debrisoquine metabolism. *Nature*, **331**, 442–6.

Gough, A. C., Miles, J. S., Spurr, N. K. *et al.* (1990). Identification of the primary gene defect at the cytochrome P450 CYP2D locus. *Nature*, **347**, 773–6.

Guengerich, F. P. (1999). Cytochrome P-450 3A4: regulation and role in drug metabolism. *Annu Rev Pharmacol Toxicol*, **39**, 1–17.

Hustert, E., Zibat, A., Presecan-Siedel, E. *et al.* (2001). Natural protein variants of pregnane X receptor with altered transactivation activity toward CYP3A4. *Drug Metab Dispos*, **29**, 1454–9.

Ingelman-Sundberg, M. (2004). Pharmacogenetics of cytochrome P450 and its applications in drug therapy: the past, present and future. *Trends Pharmacol Sci*, **25**, 193–200.

Ingelman-Sundberg, M., Oscarson, M. and McLellan, R. A. (1999). Polymorphic human cytochrome P450 enzymes: an opportunity for individualized drug treatment. *Trends Pharmacol Sci*, **20**, 342–9.

Johansson, I., Lundqvist, E., Bertilsson, L. *et al.* (1993). Inherited amplification of an active gene in the cytochrome P450 CYP2D locus as a cause of ultrarapid metabolism of debrisoquine. *Proc Natl Acad Sci USA*, **90**, 11825–9.

Johansson, I., Oscarson, M., Yue, Q. Y. *et al.* (1994). Genetic analysis of the Chinese cytochrome P4502D locus: characterization of variant CYP2D6 genes present in subjects with diminished capacity for debrisoquine hydroxylation. *Mol Pharmacol*, **46**, 452–9.

Kim, K., Johnson, J. A. and Derendorf, H. (2004). Differences in drug pharmacokinetics between East asians and Caucasians and the role of genetic polymorphisms. *J Clin Pharmacol*, **44**, 1083–105.

Kirchheiner, J., Brosen, K., Dahl, M. L. *et al.* (2001). Pharmacogenetics of antidepressants and antipsychotics: the contribution of allelic variations to the phenotype of drug response. *Acta Psychiatr Scand*, **104**, 173–92.

Kirchheiner, J., Nickchen, K., Bauer, M. *et al.* (2004). CYP2D6 and CYP2C19 genotype-based dose recommendations for antidepressants: a first step towards subpopulation-specific dosages. *Mol Psychiatry*, **9**, 442–73.

Koch, I., Weil, R., Wolbold, R., Brockmoller, J. *et al.* (2002). Interindividual variability and tissue-specificity in the expression of cytochrome P450 3A mRNA. *Drug Metab Dispos*, **30**, 1108–14.

Kuehl, P., Zhang, J., Lin, Y. *et al.* (2001). Sequence diversity in CYP3A promoters and characterization of the genetic basis of polymorphic CYP3A5 expression. *Nat Genet*, **27**, 383–91.

Lazarou, J., Pomeranz, B. H. and Corey, P. N. (1998). Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. *Jama*, **279**, 1200–5.

Lerer, B. and Macciardi, F. (2002). Pharmacogenetics of antidepressant and mood-stabilizing drugs: a review of candidate-gene studies and future research directions. *Int J Neuropsychopharmacol*, **5**, 255–75.

Lesko, L. J. and Woodcock, J. (2002). Pharmacogenomic-guided drug development: regulatory perspective. *Pharmacogenomics J*, **2**, 20–4.

Lou, Y. C., Ying, L., Bertilsson, L. and Sjoqvist, F. (1987). Low frequency of slow debrisoquine hydroxylation in a native Chinese population. *Lancet*, **2**, 852–3.

Marsh, S. and McLeod, H. L. (2004). Pharmacogenetics of irinotecan toxicity. *Pharmacogenomics*, **5**, 835–43.

Matsumura, K., Saito, T., Takahashi, Y. *et al.* (2004). Identification of a novel polymorphic enhancer of the human CYP3A4 gene. *Mol Pharmacol*, **65**, 326–34.

Nicholson, J. K. and Wilson, I. D. (2003). Opinion: understanding 'global' systems biology: metabonomics and the continuum of metabolism. *Nat Rev Drug Discov*, **2**, 668–76.

Paez, J. G., Janne, P. A., Lee, J. C. *et al.* (2004). EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science*, **304**, 1497–500.

Pao, W., Miller, V., Zakowski, M. *et al.* (2004). EGF receptor gene mutations are common in lung cancers from ''never smokers'' and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc Natl Acad Sci USA*, **101**, 13306–11.

Patel, J. and Mitra, A. K. (2001). Strategies to overcome simultaneous P-glycoprotein mediated efflux and CYP3A4 mediated metabolism of drugs. *Pharmacogenomics*, **2**, 401–15.

Pirmohamed, M., James, S., Meakin, S. *et al.* (2004). Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. *Bmj*, **329**, 15–19.

Raucy, J. L. (2003). Regulation of CYP3A4 expression in human hepatocytes by pharmaceuticals and natural products. *Drug Metab Dispos*, **31**, 533–9.

Roses, A. D. (2004). Pharmacogenetics and drug development: the path to safer and more effective drugs. *Nat Rev Genet*, **5**, 645–56.

Sachse, C., Brockmoller, J., Bauer, S. and Roots, I. (1997). Cytochrome P450 2D6 variants in a Caucasian population: allele frequencies and phenotypic consequences. *Am J Hum Genet*, **60**, 284–95.

Sackeim, H. A. (2001). The definition and meaning of treatment-resistant depression. *J Clin Psychiatry*, **62** Suppl 16, 10–17.

Schuetz, J. D., Beach, D. L. and Guzelian, P. S. (1994). Selective expression of cytochrome P450 CYP3A mRNAs in embryonic and adult human liver. *Pharmacogenetics*, **4**, 11–20.

Shimada, T., Yamazaki, H., Mimura, M., Inui, Y. and Guengerich, F. P. (1994). Interindividual variations in human liver cytochrome P-450 enzymes involved in the oxidation of drugs, carcinogens and toxic chemicals: studies with liver microsomes of 30 Japanese and 30 Caucasians. *J Pharmacol Exp Ther*, **270**, 414–23.

Smith, G., Carey, F. A., Beattie, J. *et al.* (2002). Mutations in APC, Kirsten-ras, and p53–alternative genetic pathways to colorectal cancer. *Proc Natl Acad Sci USA*, **99**, 9433–8.

Solus, J. F., Arietta, B. J., Harris, J. R. *et al.* (2004). Genetic variation in eleven phase I drug metabolism genes in an ethnically diverse population. *Pharmacogenomics*, **5**, 895–931.

Spurr, N. K., Gough, A. C., Stevenson, K. and Wolf, C. R. (1989). The human cytochrome P450 CYP3 locus: assignment to chromosome 7q22-qter. *Hum Genet*, **81**, 171–4.

Stoehlmacher, J., Park, D. J., Zhang, W. *et al.* (2002). Association between glutathione S-transferase P1, T1, and M1 genetic polymorphism and survival of patients with metastatic colorectal cancer. *J Natl Cancer Inst*, **94**, 936–42.

Stoehlmacher, J., Park, D. J., Zhang, W. *et al.* (2004). A multivariate analysis of genomic polymorphisms: prediction of clinical outcome to 5-FU/oxaliplatin combination chemotherapy in refractory colorectal cancer. *Br J Cancer*, **91**, 344–54.

Stubbins, M. J., Harries, L. W., Smith, G., Tarbit, M. H. and Wolf, C. R. (1996). Genetic analysis of the human cytochrome P450 CYP2C9 locus. *Pharmacogenetics*, **6**, 429–39.

Vastag, B. (2004). Research unveils the 'who' and 'why' of gefitinib. *J Natl Cancer Inst*, **96**, 1352–4.

Vogel, F. (1959). Moderne Problem der Humangenetik. *Ergib Inn Kinderheild*, **12**, 52–125.

Wakeling, A. E., Guy, S. P., Woodburn, J. R. *et al.* (2002). ZD1839 (Iressa): an orally active inhibitor of epidermal growth factor signaling with potential for cancer therapy. *Cancer Res*, **62**, 5749–54.

Woodahl, E. L. and Ho, R. J. (2004). The role of MDR1 genetic polymorphisms in interindividual variability in P-glycoprotein expression and function. *Curr Drug Metab*, **5**, 11–19.

Zhang, J., Kuehl, P., Green, E. D. *et al.* (2001). The human pregnane X receptor: genomic structure and identification and functional characterization of natural allelic variants. *Pharmacogenetics*, **11**, 555–72.

Index

A1555G, aminoglycoside-related hearing loss 510
Abacavir susceptibility 150
A*beta* peptide, overexpression in Alzheimer's disease 127, 128
ACE, and angiotensin converting-enzyme 366–7, 385–6
acetylators, fast and slow 170–1, 218, 520
acne vulgaris 430–1
active chromatin hub (ACH) 11
ADAM-17 440
ADAM-33 394
adiposity, as continuous variable 331–2
admixture mapping 81–2, 99, 101
  admixed population 81
adoption studies 456
  obesity 329–30
  schizophrenia 459
adrenal hyperplasia 143, 149
adrenocorticotrophic hormone, and *CYP11B1* 379, 386
adverse drug reactions, hospital admissions 516
affective disorders 454–6
  genetic epidemiology 455–6
Africa, resistance to malaria 149
African Americans, admixed population 81
aging 132–41
  familiality of exceptional longevity 134–6
    proposed multifactorial model 136–7
  genes predisposing to exceptional longevity 137–9
  relative contribution of genes, environment and luck 133
  survivors, delayers and escapers (of morbidity) 132–3
aging diseases
  age-related cataract 488–93
  age-related macular degeneration 489, 496–501
  aggregation of protein 124–6
  and new life/livestock practices 126
AGT, angiotensinogen 384–5
air pollution, asthma 391
AKT-1, V-AKT murine thymoma viral oncogene homolog-1 463

Albright hereditary osteodystrophy 333
alcoholism, depression spectrum disease 461
aldosterone synthase deficiency 380
aldosteronism
  glucocorticoid-remediable aldosteronism, and *CYP11B1* 379
  pseudohypoaldosteronism type-1 380, 382
alleles, identical by descent (IBD) 61
allelic heterogeneity 53, 100
  and new disease 53–4
*alpha*1-antitrypsin deficiency 397–8
alternative splicing 5
Alu sequence 367
Alzheimer's disease 159–60, 439–45
  A*beta* peptide overexpression 127, 128
  amyloid precursor protein 159–60, 440–1
  apolipoprotein E, *APOE* 153, 172, 441–2, 448
  genetic basis and Lewy body variant 439–40
  Huntington's disease counseling model 447–8
  other genes 444–5
  presenilin-1 and -2 442–4
  see also amyloid deposition diseases
amines, heterocyclic, and red meat 170, 217–18
amitryptyline, *CYP2D6*-dependence, dose recommendations 521
amyloid deposition diseases 118
  formation of pathological deposits 121–4
  origins 124–6
  therapeutic intervention 126–8
amyloid fibrils
  protective sequestration mechanism 125
  and serum amyloid P 127
  structure and mechanism 119–24
amyloid precursor protein and *beta*-peptide 159–60, 439–41
ancestral susceptibility models 64
anemias 316–27
  future treatment 324–5
  genetic interactions 321–2
    genotype–phenotype relationships 321–2
    modifying genes 321
  globin genes and their regulation 317–21
  prevalence 322–4
    malaria, selection of hemoglobinopathy alleles 323–4
  public health aspects 324
  see also thalassemia
Angelman syndrome 32
angiotensin converting-enzyme, and *ACE* 366–7, 385–6
angiotensinogen
  AGT 384–5
  synthesis 378
animal models
  autoimmune disease 273–4
  cardiovascular risk factors 369–71
  CFTR mutations 155
  chondrocalcinosis 417
  coronary heart disease 368–71
  familial adenomatous polyposis coli (FAP) 249–51
  hearing impairment 510, 511, 512
  inflammatory bowel disease 306
  insulin resistance 368–70
  MMR 254
  obesity 334–5
  skin color 433
  TP53 deficiency 237–8
  see also models
Ank, murine progressive ankylosis 417
ANKH 417
ankylosing spondylitis 142–3, 144, 418–20
antidepressants, *CYP2D6*-dependent, dose recommendations 521
antimorphs (dominant negative mutations) 35
antioxidants, dietary 392–3
antipsychotics, dopamine D2 receptor 462
APC 152, 248–51
  non-truncating variants 257–61
APC protein 248–51
apo A5 365
APOE 160
  apolipoprotein E, Alzheimer's disease 153, 172, 441–2, 448
apparent mineralocorticoid excess 379–81
AQP0 492
arthritis
  inflammatory 417–20
  rheumatoid 418
  spondyloarthritides 418–20
arthrogryposis multiplex congenita 204
assisted reproductive technologies (ARTs) 33
association-based gene mapping 64–72
  analysis 69–71
  consistency/strength of association 107–8
  direct/indirect association 103–5
  false discovery rate 71
  HapMap project 80, 84–9, 104
  multiple testing 71–2
  null hypothesis 70
  odds ratio (OR) 70

population attributable risk/fraction *(PAR)* 70, 93–4
power (sample size) 66–7, 95–6
sample collection and population stratification 67–8
SNP selection 68–9
stages 65
vs linkage mapping 73
see also linkage disequilibrium
asthma 391–4
environmental risk factors 173, 391–3
genetic risk factors 173, 391–3
Ataxia Telangiectasia 236
atherosclerosis, animal models 368–71
ATM 230
atopic dermatitis 431–2
atopy 392
ATP-dependent chromatin remodeling 13–14
enzymes 24
ATR-X, alpha-thalassaemia/mental retardation syndrome 26–7
Aurora-A, STK15 259
autism, specific language impairment and autism 471–80
association studies 478
autism spectrum disorders 472–3
clinical measures for autism 472–3
evidence for a genetic component to autism 472
candidate gene investigations 478–82
concordance and prevalence rates 471
cytogenetic cases 478
region of linkage on chromosome-7 479
genome screens for autism 476–8
genome screens for SLI 475–80
loci detected 477
see also speech and language disorders
autoimmune diseases 144–8, 268–76
animal models 273–4
background: genes and environment 268–70
defining non-HLA susceptibility genes 271–3
environmental effects 146
HLA (human leukocyte antigen) 270–1
immunopathology 146
mechanisms for associations 147–8
MHC 144–8
non-MHC genes 146

BACE-1, *beta*-secretase 440
Bardet–Biedl syndrome 333
Bartter's syndrome, Gitelman's variant 380
Beckwith–Wiedemann syndrome 33
beta-globin 10–11
BFSP2 492
'biobanks' 101–2
bipolar affective disorder 454–68
birth defects
disruptions, deformations, malformations 202
mortality by age groups 202
Blau syndrome, granulomas 310–11
bone
altered homeostasis
bone resorption/formation 410–13
in inflammatory arthritis 417–20
high bone mass 413–15
autosomal dominant osteopetrosis 413–15
autosomal recessive osteopetrosis 413
decreased bone resorption 413–15
pycnodysostosis 415
increased bone formation 416–17
high bone mass 416–17
juvenile Paget's disease 411, 412
low bone mass 408–10
genes contributing 409
osteoporosis 408–10
osteoporosis–pseudoglioma syndrome 410
OPG/RANK/RANKL axis 418–19
osteoblasts 406–8
osteoclasts 406–8
Paget's disease 410–12
remodeling 407
resorption and formation, increased bone turnover 411
bottleneck events 80–1
boundary elements/insulators 10
bowel cancer *see* colorectal cancer
BRCA mutations 153
breast cancer 225
carrier frequency 213, 216
smoking and cancer risk 217
breast cancer 224–32
finding high-risk individuals 154
finding low penetrance alleles 230–1
models of susceptibility 224–8
polygenic model 225–8
recessive model 228
public health implications 228–30
mammography 229–30
susceptibility alleles 228
under polygenic model 228–30
broad *beta* disease (hyperlipoproteinemia, type III) 365
bronchitis 394–400

3C-technology 10–11
C1QTNF5 498
CACC box 156
Camurati–Engelmann disease 411, 413
canalization 207
cancer
  chemotherapy 521
  and DNA methylation 29–30
  environmental exposures 213–23
  epidermal growth factor receptor (EGFR) 522
  epigenetic mechanisms 30–4
CARD15/NOD2 304, 308–11
cardiovascular risk factors
  animal models 369–71
  BMI 328
  HDL and LDL particle sizes 138
  hypertension 377–86
  and insulin resistance 367–8
carriers, Hardy–Weinberg law 62–3
case-control studies 70, 103
  controls, recruitment 177–8, 216
  genotype and environment 215–16
  vs cohort studies 103
cataract 488–93
catechol-O-methyltransferase, *COMT* 462
CCNDI, cyclin *DI* 258
CD32 286
CD36 285
  deficiency, insulin resistance 368, 369–70
CD40 285
CD209 281
CDCV hypothesis 53, 64, 75, 98, 153
Cdh23 511
  hearing loss 511
CDKN2A carriers, risk of melanoma 215, 216
cell-cycle checkpoint kinase-2, *CHEK2* 259–60
central dogma 6
centromeric heterochromatin 34
CFH, complement factor H 499–500
CFTR 155
CFTR mutations, animal models 155
CHEK2
  1100delc variant 230
  cell-cycle checkpoint kinase-2 259–60
chemotherapy, cancer 521
chi-squared contingency table test, null hypothesis 70
cholesterol
  familial hypercholesterolemia 360–4
  impaired reverse transport 364
cholesterol ester transfer protein 362
  deficiency 364
  and longevity 138
chondrocalcinosis 414, 417
chromatin 34
  active chromatin hub (ACH) 11
  active vs inactive 21
    model 21
  condensation/decondensation 34–5
  DNA packaging 11–12, 21
  epigenetic modification 20–34
  euchromatin 35
  loops 16
  modification and transcription 11–14
  remodeling 11–15, 23–4
  status within genome 15
  structure 11
chromosomal architecture 4
chromosomal disorders 202–3
chromosome territories 16
chronic obstructive pulmonary disease (COPD) 394–400
  candidate genes 399
  environmental risk factors 396
  smoking-related 394–6
chylomicrons 362
cis-acting regulatory elements 4, 8–10, 35
CJD *see* Creutzfeldt–Jakob disease
CLCNKB 380
clinical research 182–7
  case-control studies 70, 103, 215–16
  controls, recruitment 177–8, 216
  developing world 187–8
  ethical oversight 183–5
  ethico-legal issues in genetics 182–8
  family studies 178–80
  feedback for disease prevention 180–1
  informed consent 185–6
  instruments 182–3
  publication 187
  randomized controlled trials 186–7
  remuneration 186
  risk–benefit analysis 184–5
clusterin 124
coactivator proteins, transcriptional activation 9, 11
coalescent process 49
cochlear implant 508
codons, premature termination 7

cohort studies, genotype and environment 214–15
COL1A1 408
  collagens 162
colorectal cancer 245–67
  APC 152, 248–51
  common genetic variation and susceptibility 257–61
    non-truncating variants in *APC* 257–61
    other putative risk alleles 260–1
  dietary aspects 217–18, 217–19
  epidemiology 245–7
  familial adenomatous polyposis coli (FAP) 247–51
  folate and 219–20
  genetic susceptibility 247
  hereditary non-polyposis colorectal cancer (HNPCC) 251–5
  MYH protein 256–7
  NAT2, fast and slow acetylators 170–1, 218, 520
  rare dominant disorders 255–6
  recessive inheritance 256–7
  smoking 219
common disease–common variant (CDCV) hypothesis 53, 64, 75, 98, 153
complement factor H, *CFH* 499–500
complex disease
  genetic architecture 97–100
  models 52–3
  trait mapping 92–112
    characteristics of populations 100–2
    detection of rare variants 99
    expression of QTLs 105–6
    family-based/population-based approaches 98, 102–5
    genetic effect sizes 97–100
    heritability estimates 94
    preliminary steps 94–7
    results 107–9
COMT, catechol-O-methyltransferase 462
congenital adrenal hyperplasia 143, 149
  hypertension syndromes 380
congenital heart disease, genes identified 203
coronary heart disease 359–76
  environmental contribution 359–60
    risk factors 360
  evidence for genetic basis 360
  genome-wide studies 365–8
  incidence 359
  monogenic and complex dyslipidemias 361–5
  rodent models 368–71
cortisol, and apparent mineralocorticoid excess 379–81
counseling
  genetic disease 447–8
  legal-ethical issues 190–1
Cowden disease (CD) 255–6
CpG island 26, 35, 253
CpGs, methylated 25–6
craniometaphyseal dysplasia 414, 417
Creutzfeldt–Jakob disease
  genes 284
  regulation of prion protein by quinacrine 126–7
Crohn's disease *see* inflammatory bowel disease
alpha-crystallin, *CRY* genes 491–2
CTLA4, immunomodulatory co-receptor 147
CX46 and *CX50* 492
cyclin *DI*, *CCNDI* 258
CYP2C, *CYP2D*, and *CYP3A*, drug metabolism 519–20
CYP2D6
  CYP2D6-dependent antidepressants, dose recommendations 521
  ethnic diversity 523
  ultrarapid metaboliser 519
CYP11B1
  and adrenocorticotrophic hormone 379, 386
  congenital adrenal hyperplasia 380
CYP11B1 and *CYP11B2* 386
  glucocorticoid-remediable aldosteronism 379–81
CYP11B2, aldosterone synthase deficiency 379–81
CYP17, congenital adrenal hyperplasia 380
CYP21, and 21-hydroxylase deficiency 149, 380
cystic fibrosis 154–5
Cystic Fibrosis Genotype–Phenotype Consortium 155
cytochrome *P-450* 259
cytoskeletal proteins 492

DAO (D-amino-acid-oxidase), G72 461–2
DC-SIGN 281, 287
deafness, *GJB2* 509
deformations and malformations 202
dementia
  familial British dementia (FBD) 445
  familial encephalopathy, neuronal dementia with neuroserpin deposits 445
  familial non-specific dementia 445
  fronto-temporal lobe dementia (FTD) 445–6
  non-specific dementia 445
  pharmacogenomics 448–9
  predictive genetic testing 447–8
  tau mutations in FTDP-17 446–7
  see also Alzheimer's disease

dengue virus, associated genes 284
depression
  antidepressants, *CYP2D6*-dependent, dose recommendations 521
  and *MTHFR* 460–1
  unipolar/bipolar 454–6
    twin studies 455
depression spectrum disease, and alcoholism 461
'detectance' 94
  improving 95
developing world, clinical research 187–8
developmental disorders 201–12
  environmental contribution 204, 207–9
  etiological groups 202–4
  genetic contribution 204–7
  mortality by age groups 202
  prevalence 201–2
diabetes mellitus type-1
  HLA-DQ allele 144, 145, 270–1
  pregnancy
    structural malformations 208
    vitamins 209
diabetes mellitus type-2 344–58
  and BMI 328
  candidate genes 351–3
  definitions 344
  environmental factors 167–8, 346–8
  genes 345–6, 348
  heritability of quantitative traits 346
  medical care 353
  and metabolic syndrome 367–8
  next global epidemic 344–5
  susceptibility gene identification 348–51
    genome scans for linkage in multifactorial T2D 350–1
    maturity onset diabetes of the young (MODY) 348–50
    other monogenic syndromes 350
  thrifty genes and selective insulin resistance 167–8
diet
  antioxidants 392–3, 397
  and asthma 392
  and colorectal cancer 217–19
  red meat, heterocyclic amines 170, 217–18
  and respiratory function 397
disability
  legal-ethical issues 191–3
  postponement in aged people 132
disability-adjusted life-years (DALYs) 278
DISC1 and *DISC2* 462
disruptions, deformations, malformations 202
DNA deletions, disorders 203
DNA methylation 15, 35
  and cancer 29–30, 219
  transcriptional repression 25–6
  see also folate
DNA mismatch repair genes 251–4
  DNA instability in tumors 253
DNA modification 24–5
dominant negative mutations 35
dopamine receptor, *DRD3* 462–3
DRD3, dopamine receptor 462–3
Drosophila, *DSCAM* 6
drug responses *see* pharmacogenomics
DTNBP1, dysbindin 461
Duffy antigen/chemokine receptor 285
dysbetalipoproteinemia 365
dysbindin, *DTNBP1* 461
dyslipidemia, animal models 368–71

embryonic screening, and selection, legal-ethical 189–90
emphysema 394–400
  and *alpha*1-antitrypsin deficiency 397–8
ENaC, epithelial sodium channel 382–3
encephalopathies, transmissible, aggregation of protein 124
endophenotypes 97
endoplasmic reticulum, protein folding 116
endosteosis, Van Buchem's disease 414, 416
enhancer elements 9, 10–11, 35
environment 164–75
  definitions and scope 164–5
  as determinant of genotype and disease 166–7
  estimations of health impact 165–6
  gene–environment interactions 170–1
  major mendelian genes vs minor polygenes as predisposers 171–2
  prevention strategies: whole populations, high-risk groups or selected individuals 172–3
environmental exposures
  autoimmune diseases 146
  and cancer 213–23
  developmental disorders 204, 207–9
epidermal growth factor receptor (EGFR), cancer 522, 523
epigenetic mechanisms 14–15, 20–34
  in cancer 30–4
  defined 14, 20, 35
  malformations 203
epistasis, gene–gene interactions 106
epithelial sodium channel, ENaC 382–3

equality of access, and discrimination 193–4
ERBB2, tyrosin receptor 522
ethico-legal issues in genetics 176–98
  clinical research settings 182–8
  ethical oversight in clinical research 183–5
  feedback for disease prevention 180–1
  legal-ethical issues in clinical settings 188–94
    attitudes and approaches to disability 191–3
    embryonic screening (PGD) and selection 189–90
    equality of access and discrimination 193–4
    risks of coercion in counseling 190–1
  legal-ethical issues in research settings 176–7
  ownership and governance issues 181–2
  pharmacogenomics 526
  recruitment for case-control studies 177–8
  recruitment for family studies 178–80
Ewens sampling theory 48–9
exonic splicing enhancers 157–9
exons 4, 6
  alternative 6
expected likelihood of odds (eLOD) score 76
expression quantitative trait loci (eQTLs) 105–6
expressivity, single gene disorders 153–5
'extended haplotype homozygosity' (EHH) 84, 89
extracellular matrix (ECM) proteins 497
extreme discordant and concordant (EDAC) sampling 75, 102
eye diseases 488–504

Factor VII 367
familial adenomatous polyposis coli (FAP) 247–51
  animal models 249–51
  cellular functions of APC protein 248–9
  FAP germline mutations 248
  FAP modifiers 249–51
familial British dementia (FBD) 445
familial combined hyperlipidemia (FCHL) 364–5
familial defective apolipoprotein B-100 (FDB) 361–4
familial encephalopathy, with neuronal dementia and neuroserpin deposits 445
familial expansile osteolysis 411–13
familial hypercholesterolemia 360–4
  frequency 360
  mortality 161
  model modification of phenotype 161
  and statins 160–1
familial non-specific dementia 445
family studies
  clinical research 178–80
  major design and size constraints 216–17
FEN1B 445
fetal neuromuscular proteins, arthrogryposis multiplex congenita 204
fibulin1-6 497
fitness, defined 49
FLAP, and 5-lipoxygenase activating protein 366
folate deficiency
  and colorectal cancer 219–20
  and neural tube defects (NTDs) 207–8
founder populations, inbred 101
FOXC1 158
FOXP2 469
  genomic structure and function 474
fragile X syndrome 333
fronto-temporal lobe dementia (FTD) 445–6
FSHD 158
F_{ST} statistic 89

G protein signalling-4, *RGS4* 462
G6PD 285
G72, DAO (D-amino-acid-oxidase) 461–2
gain-of-function, atherosclerosis, animal models 369, 370
galectin-2, and myocardial infarction 367
GALK1 492
Gardner's syndrome 248
GCD 349
gefitinib 522
gene deletion syndromes 203
gene expression 6–19
gene mapping 59–79
  association-based gene mapping 64–72
  future prospects and impact on health 76–7
  homozygosity mapping 63
  linkage mapping in families 60–4
  null hypothesis 70
  quantitative trait loci (QTLs) 54–5, 72–5
gene silencing (transcriptional repression) 37
gene therapy, *TP53* deficiency 241
genes
  defined 3
  number of human genes 5, 20
  number of proteins 5
  silenced 11
genetic counseling 190–1, 447–8
genetic diagnosis, preimplantation (PGD) 189–94
genetic disease, incidence 45

genetic drift 46–8
genetic testing
  permutation testing 71–2
  pharmacogenomics 448–9
  predictive genetic testing 447–8
genetic variation
  classification 108
  population, measurement 45–6
  see also single nucleotide polymorphisms (SNPs)
genomic disorders 203
genomic imprinting 35
  and ARTs 33–4
  diseases of 31–4
  model 32
genotype relative risk (GRR) 67, 75
Gitelman's variant, Bartter's syndrome 380
GJB2, deafness 509
glaucoma 493–6
  primary angle closure 495–6
  primary open angle 493–5
GLC1A-G 494
GLI3 158
globin genes and their regulation 317–21
  alpha globin cluster 317–18
  alpha thalassemia 318–19
    mutations in *trans*-acting factor 319
  beta globin cluster 318, 319
glucocorticoid-remediable aldosteronism, *CYP11B1* and *CYP11B2* 379–81
glucokinase 349
glucose-6-phosphate dehydrogenase, *G6PD* 285
glutamate receptor metabotropic-3, *GRM3* 463
glutathione-S-transferase 399, 521
glycophorin C 285
Gordon's syndrome, pseudohypoaldosteronism type-1 380, 382
granulomas, Blau syndrome 310–11
GRM3, glutamate receptor metabotropic-3 463
growth control, SWI/SNF complex 24

HapMap project
  association-based gene mapping 80, 84–9, 104
  fine localization 88–9
  identifying evidence for selection 89
  SCN8a dataset 86–8
  tagging SNPs 84–6
haptoglobin 284
Hardy–Weinberg law 46–8
  carriers 62–3
HBA 158
HBB 158
HDL cholesterol, in premature CHD 364
HDL lipoproteins, metabolism 362
hearing impairment 505–27
  age-related 509–11
  aminoglycoside-related 510
  environmental causes, aminoglycoside-related 510
  genetic heterogeneity 511
  mechanisms 505–7
  mendelian disorders 511
  prostheses 507–8
heart disease
  genes identified in congenital syndromes 203
  see also myocardial infarction
heat shock proteins (molecular chaperones) 113–15
Helicobacter pylori, associated genes 283
heme-oxygenase-1, *HMOX-1* 398–9
hemochromatosis 143, 149
hemoglobin
  HbF 321
  structure 316
hemoglobinopathies
  geographical distribution 322
  and malaria 323–4
  public health aspects 324
hemolytic anemia 285
hepatitis virus B and C, associated genes 284
hepatocyte nuclear factors
  1-*beta* 349
  4-*alpha* 349
hereditary non-polyposis colorectal cancer (HNPCC: Lynch syndrome) 154, 251–5
  DNA mismatch repair genes 251–4
  MMR defects and DNA instability in tumors 253
  MMR mutation spectrum and genotype-phenotype correlations 252–3
  molecular mechanisms of colorectal carcinogenesis and defective MMR 253–4
  other candidate susceptibility genes 254–5
heritability estimates, trait mapping 94
heterochromatin 11, 21, 35
  centromeric 34
heterochromatin protein-1 (HP1) (nuclear factor) 22–3
heterozygosity, genetic markers 61
hexokinase-4 349
Hirschsprung disease 205–6
  sibling recurrence risk 205
histone acetyltransferases (HATs) 12–13, 22, 31

histone deacetylases (HDACs) 22–4
histone methyltransferases (HMTs) 23
histones 11, 21
  histone code 14
  modification 22–3
HIV infection and AIDS 290–1
  genes 283–4
  and MHC system 148–9
HLA (human leukocyte antigen) 270–1, 279
  HLA-B27 418–20
  rheumatoid arthritis 418
  tuberculosis 287
  see also MHC
HMG proteins 11
HMOX-1, heme-oxygenase-1 398
HNF1A 349
HNF4A 349
homozygosity mapping 63
hormone replacement therapy, and colorectal cancer 245
hospital admissions, adverse drug reactions 516
Hoxd complex 158
HP1, heterochromatin protein-1 22–3
HSD11B2, 11-*beta* hydroxysteroid dehydrogenase type-2 381
HSF4 492
5-HT2A, serotonin receptor 462–3
Human Genome Project, media claims *xviii–xiv*
human subjects research *see* clinical research
human *Toll*-like receptors 272, 279–81
  reported ligands 280
Huntington's disease 154
  counseling 447–8
11-*beta* hydroxylase
  CYP11B1, glucocorticoid-remediable aldosteronism 379
  deficiency, and CAH 380
17-*alpha*-hydroxylase
  deficiency, and CAH 380
  hypercholesterolemia 364
21-hydroxylase deficiency, and *CYP21* 149, 380
11-*beta* hydroxysteroid dehydrogenase type-2, *HSD11B2* 381
hygiene hypothesis 392
hypercholesterolemia, monogenic disorders 360–4
hyperlipidemia, familial combined (FCHL) 364–5
hyperlipoproteinemia, type III 365
hypertension 377–90
  defined 377
  HyperGen study 384
  intermediate phenotypes 385–6
  linkage studies 383–5
  mendelian forms 378–83
  undertreatment 377
hypertension and hypokalemic alkalosis syndrome 381
hypokalemic alkalosis syndrome 381

ICAM Kilifi 285
identical by descent (IBD) alleles 61, 74
IFNAR1, interferon-*alpha* receptor-1 286, 289
infectious diseases 277–301
  deaths (1999) 278
  disability-adjusted life-years (DALYs) 278
  historical perspective 277–8
  human *Toll*-like receptors, reported ligands 280
  and immature immune system 146
  malaria 281–6
  MHC 148–9
  mycobacterial disease 286–9
  Pseudomonas infections 149
  reportedly associated genes 282–4
  twin and adoption studies 278–81
infertility, CF patients 155
inflammatory arthritis 417–20
inflammatory bowel disease 302–15
  animal models developing a colitis 306
  Crohn's, and *NOD2* 53, 75
  environmental factors 303–4
    gut flora hypothesis 304
  genetic factors 305–11
  and *NOD2* 53, 75
  positional cloning of IBD genes 307–8
    functional candidate genes investigated in IBD 307
    localization of IBD loci 308
  twin studies from Sweden and Denmark 306
informed consent 185–6
inositol-5-phosphatase, *SHIP2* 385
insulators 10
insulin degrading enzyme 440
insulin promotor factor-1 349
insulin recptor, mutations 350
insulin resistance
  CD36 deficiency 368, 369–70
  and CV risk factors 367–8
  selective, diabetes mellitus type-2 167–8
insulin-like growth factor-1 260
interferon-*alpha* receptor-1, *IFNAR1* 286, 289
interferon-*gamma* 288
international collaboration 109
introns 4

IPF1 349
Iressa 521
iris, congenital absence, *PAX6* 157

juvenile polyposis syndrome (JPS) 255

KIR, MHC-associated disease 147

LCT 158
LDL cholesterol, hypercholesterolemia 160–1, 363
LDL receptor
  animal models 369
  and CHD risk 153
  LRP5 408
  LRP5 mutations, familial hypercholesterolemia 361
leishmaniasis 282, 287
LEKT1 394
leprosy 288–9
  associated genes 283
leptin, reponse to leptin therapy 335
leptin deficiency 334–5
  partial, in heterozygotes 335
leptin receptor deficiency 335–6
leukotriene A4 hydrolase, and myocardial infarction 367
Lewy bodies 121
Liddle's syndrome 380, 382–3
LIM2 492
linkage disequilibrium 51–2, 59–60
  admixture mapping 81–2
  'blocks' 83
  estimation of LD 66
  genome-wide patterning 80–1
  localized LD 82–9
    and recombination 82–3
    selection and shaping 83–5
  MALD method 82
  shared IBDs 104
  signal of selection 8
  see also association
linkage mapping
  in families 60–4
  power requirements 102
  required no. of sibling pairs to detect 74
  transmission disequilibrium test (TDT) 72
  vs association-based gene mapping 73
linkage regions 63
5-lipoxygenase activating protein, and *FLAP* 366
lipoproteins, metabolism 362
locus control regions 9–10, 157
Log of odds (LOD) score 63, 73, 76–7
  expected (ELOD) 76
long distance regulators 156–7
longevity studies 132–41
loss-of-function, atherosclerosis, animal models 369, 370
LRP5
  LDL receptor 408, 416
  mutations, familial hypercholesterolemia 361
lung disease, environmental risk factors 396
lymphotoxin-*alpha* 367
lysine histone acetyltransferases (HATs) 22
lysine histone methyltransferases (HMTs) 23
lysozyme, and aging diseases 126

macular degeneration 489, 496–501
MAF 158, 492
malaria 281–6
  associated genes 282
  and hemoglobinopathies 323–4
  resistance 149, 285–6
  sickle-cell trait and disease 281, 323
  susceptibility, MHC 286
  and thalassemia 317
MALD method, linkage disequilibrium 82
malformations 202
mammography 229–30
maternally inherited diabetes and deafness (MIDD) 350
matrix attachment sites (MARs) 10
maturity onset diabetes of the young (MODY) 348–50
MC1R
  and *p16*, human pigmentation 435
  and skin cancer 432–4
  skin color and melanoma risk 220–1, 433
MC4R, melanocortin 4 receptor 337
MDM2 236
MED1 253
melanin 220–1, 432–3
  pheomelanin 221, 432–3
melanocortin 4 receptor, *MC4R* deficiency 337
melanoma risk
  in *CDKN2A* carriers 215, 216
  skin color 220
    MC1R 220–1
MELAS syndrome 350
membrane transport proteins 492
membranoproliferative glomerulonephritis, *CFH* mutations 500

mendelian randomization 106–7
mental retardation (ATR-X syndrome) 26–7
MEOX1 416
metabolic syndrome 367
metabonomics 518
methyl tetrahydrofolate reductase, *MTHFR* 258–9, 409
methylated CpGs 25–6
  methyl-CpG binding protein (MeCP2) 27–8, 30
methylenetetrahydrofolate reductase, C677T variant 219
MHC 142–51
  Abacavir susceptibility 150
  autoimmune diseases 144–8
  features/general considerations 142–4
  infectious diseases 144, 148–9
    malaria susceptibility 286
    Pseudomonas infections 149
  linkage disequilibrium across region 144
  psoriasis 149–50
  TAP transporters 149
microarray analysis 35–6
microbial burden, reduction, and immature immune system 146
microsatellite instability 219, 253
microsatellite markers 61
microsomal epoxide hydrolase, *mEPHX* 399
mineralocorticoid receptor *(MR)* 382
mismatch repair genes 251–4
  mutations 252
mismatch repair proteins, hMSH2 and hMLH1 220
mitochondrial genome, *A1555G*, and aminoglycoside-related hearing loss 510
MLH1, *MLH3* 251–3
models
  active vs inactive chromatin 21
  ancestral susceptibility 64
  complex disease 52–3
  genomic imprinting 32
  multi-locus 51–2
  single-locus natural selection 49–51
  Wright–Fisher model 47
  see also animal models
modifier loci 155–6
molecular chaperones
  clusterin 124
  functions 125
  protein folding 113–15
molecular mimicry hypothesis, T cells 147
Mormon families, linkage in schizophrenia 460
mRNAs
  from single gene 5
  structure variations 5–6
MSH6 251–2
MTHFR
  and depression 460–1
  methyl tetrahydrofolate reductase 258–9, 409
Muckle–Wells syndrome, *PYPAF1* 311
multiple sclerosis 168–70
  ultra-violet radiation (UVR) 169
murine thymoma viral oncogene homolog-1, *V-AKT* 463
mutation–selection balance 50
mutations
  de novo, newborns 45
  dominant negative 35
  genomic mutation rate 45
  incidence, Mendelian 46
  rates, new mutations 60
mycobacterial disease 286–9
  associated genes 283
  leprosy 288–9
  rare atypical infections 289
  tuberculosis 286–8
myeloid/lymphoid or mixed lineage leukemia 31
MYH protein
  -associated colorectal neoplasia 256–7
  function and interactions 256–7
MYO7A, Usher syndrome 511
myocardial infarction
  ACE 385
  and galectin-2 367
  and leukotriene A4 hydrolase 367
myocilin, *MYOC* 494–5

N-acetyl transferase 218, 520
NAT2, fast and slow acetylators 520
  colorectal cancer 170–1, 218
natural resistance-associated macrophage protein gene, *NRAMP1* 287
natural selection, model 49–51
neonatal jaundice 285
neprilysin 440
Netherton's syndrome, *SPINK5* 394, 431–2
neural tube defects (NTDs), and folic acid 207–8
neuregulin-1, *NRG-1* 461
NEUROD1 349
neurogenic differentiation-1 349
neuronal dementia, with neuroserpin deposits 445
neutral apoptosis-regulated convertase-1 (NARC-1) 361–2

nicastrin 443
NICD, Notch intracellular domain 443
nitric acid synthase 2A, *NOS2A* 286
NOD and *NOD2* 53, 75, 272
  Crohn's disease 53, 75
nonsense-mediated RNA decay (NMD) 7
Norwalk virus, associated genes 284
NOS2A 286
Notch intracellular domain, *NICD* 443
NRAMP1 287
NSAIDS, and colorectal cancer 246
nuclear compartmentalization 15–16
nuclear factor
  heterochromatin protein-1 (HP1) 22–3
  kappa B, RANKL 406, 410–12, 418
nuclear matrix 36
nucleolus 16
nucleosomes 21, 22, 36
null hypothesis, chi-squared contingency table test 70
number needed to screen (NNS) 228
number needed to treat (NNT), risk analysis 228
nutritional defects, teratogenesis 207–8

obesity 328–43
  adiposity measurement 331
  adoption studies 329–30
  animal models 334–5
  common obesity
    association studies 339
    linkage studies 337–9
  determining susceptibility 329–31
  gene/environment interaction 169–70, 339–40
  heritability of intermediate traits 330
  maternal, as risk factor for disease 209
  methodological issues 331
  monogenic disorders leading to human obesity 332
  neuronal pathwayays involved in regulation of body weight 334
  novel human monogenic obesity syndromes 334
  progress in identification of human obesity genes 332–9
  reponse to leptin therapy 335
  threat to public health 328–9
  twin studies 329, 330
obstructive pulmonary disease 391–405
  candidate genes 399
  environmental risk factors 396
  smoking-related 394–6
odds ratio (OR), association-based gene mapping 70
OPG/RANK/RANKL axis, bone destruction 418–19
ophthalmic diseases 488–93
optineurin 495
osteoblasts 406–8
osteoclasts 406–8
osteolysis, familial expansile 411, 412–13
osteopetrosis
  autosomal dominant 413–15
  autosomal recessive 413–14
  with renal tubular acidosis 414–15
osteoporosis 408–10
osteoporosis–pseudoglioma syndrome 410

P2X7 purinergic receptor 287
p16, human pigmentation 435
p53
  cell-cycle control, apoptosis and DNA repair 236–7
  DeltaNp53 isoform 237
  effectors of p53 functions 238
  negative regulation 237
  normal suppression of proliferation 237–9
  structure 235
p53 and cancer 233, 240–2
  experimental therapy using p53 as target 241–2
  p53 in normal suppression of proliferation 237–9
  see also TP53
p53/Mdm2 236
P450, and *CYP3A4* 519–20
Paget's disease of bone 410–12
PAIRED box 5
parasitic elements 36
parathyroid hormone receptor type-1, *PTHR1* 410
PASG 27
pathogen-associated molecular patterns (PAMPs) 279
PAX6 158
  iris, congenital absence 157
  transcription factor, alternative splicing 5
PCR, real time 36
PCSK9, hypercholesterolemia 361–2
PEN-2 443
penetrance, single gene disorders 153–5
peptides and proteins, aggregation rates, effect of mutations 122
peptidylprolyl isomerases 115
permutation testing 71–2
Peutz–Jeghers syndrome (PJS) 255
pharmacogenomics 516–27
  clinical applications 520–2
  current applications 524–5

CYP3A, CYP2D, CYP2C families 519–20
databases 517
defined 516
ethico-legal issues 526
ethnic differences 522–3
future 525
genetic testing 448–9
metabonomics 518
phenotype informativeness, improving 'detectance' 94–5
phenotype–genotype relationships 96
pheomelanins 221, 432–3
phylogenetic footprinting 17
PITX2 158
PITX3 492
plasminogen activator inhibitor-1 368
Plasmodium spp. 282
resistance loci 286
polyadenylation sites 5
polypeptides, synthesis on ribosome 123
population attributable risks/fractions *(PAR)* 70, 93–4
population genetics 44–58
admixture mapping 81–2
experimental results 54–5
genetic drift 46–8
Hardy–Weinberg law 46–8
measurement of variation 45–6
suitability of populations 100–1
position effect variegation 22, 36
POU3F4 158
power (sample size), association-based gene mapping 66–7, 95–6
Prader–Willi syndrome 32–3, 332–3
preaxial polydactyly (PPD) 157
predictive genetic testing 447–8
preimplantation genetic diagnosis (PGD) 189–94
premature termination codons (PTCs) 7
presenilin-1, *PS1* 442–3
presenilin-2, *PS2* 443–4
prion diseases 277
genes 284
prion protein, regulation by quinacrine 126–7
pro-opiomelanocortin deficiency 336
progesterone-sensitive hypertension 380
prohormone convertase 1 deficiency 336–7
Proliferation-associated SNF2-like gene (PASG) 27
proline dehydrogenase, *PRODH* 462
promotor region, hypermethylation, *RB* 30
promotors 36
and transcription machinery 7–8
protein aggregation, transmissible encephalopathies 124
protein disulphide isomerases 115
protein folding
catalysts 115
molecular chaperones 113–15
protein misfolding 116–17
diseases 117–18
proteome 4–6
PS1, presenilin-1 442–4
PS2, presenilin-2 443–4
pseudohypoaldosteronism type-1 380, 382
Pseudomonas infections, and MHC 149
psoriasis 149–50, 429–30
psychiatric disorders 454–68
adoption studies 459
genetic epidemiology 455–6
linkage analysis 458–60
positional/functional candidate genes 461–3
PTCs, premature termination codons 7
PTEN
breast cancer 225
Cowden disease 255–6
juvenile polyposis 255
PTHR1 410
pycnodysostosis 414–15
PYPAF1, Muckle–Wells syndrome 311

quantitative trait loci (QTLs) 54–5, 72–5
expression 105–6
linkage mapping 337–8
linkage mapping vs association approaches 73
quantitative traits, complex disease 97
quinacrine, regulation of prion protein 126–7

randomized controlled trials 186–7
RANKL, nuclear factor kappa B 406, 410–12, 418
recombination, and linkage disequilibrium 51–2
red meat, heterocyclic amines 217–18
regulated unproductive splicing and translation (RUST) 7
renal tubular acidosis 415
renin–angiotensin–aldosterone system 378–9
replication timing 36
repressors 9
research in complex disease 92–112
respiratory syncytial virus, associated genes 284
restriction enzymes 36
RET, Hirschsprung disease 205–6

retinitis pigmentosa 153
retinoblastoma gene, promotor region hypermethylation 30
Rett syndrome 27–9
RGS4, G protein signalling-4 462
rheumatoid arthritis 417–18
ribosome, synthesis of polypeptide chain 123
risk analysis
genotype and environment 214–15
number needed to screen (NNS) 228
number needed to treat (NNT) 228
susceptibility genes 228–30
risk profiles 229
risk–benefit analysis, clinical research 184–5
RNA, micro-RNA 159
RNA polymerases (RNAPs) 6–7, 8
RNA Pol II 36
RNA splicing 4
RNA-TRAP 10
RNAi (interference) 36–7
rodent models *see* animal models

Salmonella typhi, associated genes 283
salt and water balance, renin–angiotensin–aldosterone system 378–9
sampling
Ewens theory 48–9
extremely discordant and concordant (EDAC) 75
increasing sample size 104
power, association-based gene mapping 66–7, 95–6
scaffold attachment sites 10
schistosomiasis 282
schizophrenia 454–61
adoption data 459
functional candidate genes 462–3
linkage analysis 458–9
positional candidate genes 461–2
sib-pair analysis 458
twin studies 457
sclerosteosis, and Van Buchem's disease 416
sclerostin, *SOST* 416
SCN8a dataset, HapMap project 86–8
beta-secretase, BACE-1 440
seeding phenomena, transmissible encephalopathies 124
sepsis/septic shock, associated genes 283
serotonin receptor, *5-HT2A* 462
serpin deposits 445
Seventh Day Adventists, health and life expectancy 133
SHFM1 158
SHH 157, 158
SHIP2, inositol-5-phosphatase 385
sib-pair analysis 458
sibling risk ratio (*lambda*-s) 143, 398
sickle-cell trait and disease, malaria 281, 323
signal of selection, linkage disequilibrium 8
single gene disorders 152–63
estimated number 152
genetic risk factors 161–2
genome sequence pinpoints novel potential regulatory variants 159
identifying modifier loci 155–6
locus and allelic heterogeneity 153
mutational spectra 156–9
genome sequence, potential regulatory variants 159
long distance regulators 156–7
new mutation target: exonic splicing enhancers 157–9
plethora of functional mutations 159
position-effect genes in human diseases 158
novel therapeutic approaches 159–60
penetrance, expressivity and genetic modifiers 153–5
single nucleotide polymorphisms (SNPs) 59–60
case–control studies 70
haplotype tagging SNPs 83, 84
HapMap project 80, 84–9
markers for linkage 64
selection, association-based gene mapping 68–9
tSNP design 86–8
siRNA (small interference RNA) 37
sitosterolemia 363, 364
SIX3 158
skeletal disorders 406–26
altered bone homeostasis 417–20
altered homeostasis of bone resorption and bone formation 410–13
bone remodeling 407
high bone mass 413–15
increased bone formation 416–17
low bone mass 408–10
skin cancers 432–5
human pigmentation 432–5
MC1R and *p16* 435
melanoma risk
CDKN2A carriers 215, 216
skin color, *MC1R* 220–1
skin color
animal models 433
constitutive/facultative 433
melanoma risk 220–1

skin diseases 427–38
acne (acne vulgaris) 430–1
atopic dermatitis 431–2
psoriasis 429–30
spatiotemporal patterns of common skin diseases 428
SMAD, TGF*beta*/SMAD pathway 254–5
smoking
cancer risk 173, 217–18
BRCA mutations 217
chronic obstructive pulmonary disease (COPD) 394–6, 396–9
colorectal cancer 219
pack years, defined 395
sonic hedgehog (SHH) 157
Sorsby fundus dystrophy 499
SOST, sclerostin 416
SOX9 158
SPCH1 locus
SPCH1–FOXP2 473–4
specific language impairment 473–4
speech and language disorders 469–87
epidemiology 469–73
glossary 482–3
clinical terms 482–3
technical terms 483
molecular genetic approaches 473–82
expression studies of *FOXP2* 474–5
identification of *SPCH1–FOXP2* 473–4
proposed molecular mechanisms 475
monogenic speech and language disorder 470
KE family phenotype/pedigree 470
patient *CS* 470
specific language impairment
and autism 471–80
autism spectrum disorders 472–3
clinical measures for SLI 472
evidence fora genetic component to SLI 471–2
genome screens for SLI 475–80
SPCH1 473–4
see also autism, specific language impairment
SPINK5, Netherton's syndrome 394, 431–2
spondyloarthritides 418–20
SRY 158
statins, and familial hypercholesterolemia 160–1
STK15, Aurora-A 259
surfactant protein A 288
SWI/SNF complex, growth control 24
alpha-synuclein 127
T cells
MHC class I peptide presentation 142
molecular mimicry hypothesis 147
shaping, and microbial burden reduction 146
TACE 440
TATA-binding protein 8
tau mutations in FTDP-17 446–7
teratogenesis 204, 207–9
in maternal diabetes mellitus type 1 208
nutritional defects 207–8
TGFB1 407, 413
alpha-thalassemia/mental retardation syndrome (ATR-X) 26–7
thalassemia 156, 317, 318–21
alpha thalassemia
mutations of *alpha* globin 318–19
mutations in *trans*-acting factor 319
alphabeta thalassemia 321
beta thalassemia
mutations in *beta* globin 156, 320
mutations in *trans*-acting factors 320–1
geographical distribution 322
long distance regulators 156–7
and malaria 317
public health aspects 324
thalidomide 207
thrifty genes, and selective insulin resistance 167–8
thrifty phenotype hypothesis 347
TIMP-3 498
TNFRSF1B, tumor necrosis factor receptor family 1B 409
TNFRSF11A 412
Toll-like receptors 272, 279–81
reported ligands and selected references 280
TP53 233–44
breast cancer 225
characteristics 233–5
deficiency, gene therapy 241
frequency of mutations in cancer 239–40
beyond loss of suppression: gain-of-function effects 240
detection of mutations 241
mutagenesis vs selection 239–40
gene structure 234
master gene in normal and tumor suppression 233–44
p53 and cancer 235, 240–2
regulation and function 236–9
regulation of suppression 236
TP53 family 233–5
see also p53 and cancer

TP63 234
TP73 234
trans-acting transcription factors 4, 35
transcription 6–7
  termination 7
transcription factors 7–9, 492
  general (GTFs) 7–8
  trans-acting 4, 35
transcriptional activation 9, 11
  coactivator proteins 9, 11
transcriptional regulation 7–15
  boundary elements/insulators 10
  and chromatin remodeling 11–15
  cis-regulatory elements 4, 8–10
  and disease 16–17
  enhancer action 8–9, 10–11
  promotors and transcription machinery 7–8
  RNA polymerases (RNAPs) 8
transcriptional repression 37
  by DNA methylation 25–6
transcriptome 4–6
transforming growth factor
  TGF*beta*/SMAD pathway 254–5
  TGF*beta*1 407, 413
transmissible encephalopathies
  aggregation of protein 124
  seeding phenomena 124
transmission disequilibrium test (TDT) 72, 102, 143
transposons 37
transthyretin, in amyloid disease 126
trichothiodystrophy 320
trisomies 203
tuberculosis 282, 286–8
  associated genes 282, 287–8
  and vitamin D 287
tumor necrosis factor *alpha* 286, 399
tumor necrosis factor receptor family 1B, *TNFRSF1B* 409
twin studies
  and adoption studies, infectious diseases 278–81
  asthma 173
  atopic dermatitis 431–2
  depression (UPD and BPD) 455
  inflammatory bowel disease 306
  obesity 329, 330
  schizophrenia 457
  sibling risk ratio 143
TWIST 158
two-hybrid screening 37
typhoid, associated genes 283
tyrosine kinase receptor, *ERBB2* 522

ubiquitin–proteasome system 124
UDP-glucuronyl transferase 1A1 521
ulcerative colitis *see* inflammatory bowel disease
ultra-violet radiation (UVR) 169
  and MS 169
ultrarapid metaboliser, and *CYP2D6* 519
unipolar depression 454–68
Usher syndrome, *MYO7A* 511

V-AKT murine thymoma viral oncogene homolog-1, *AKT-1* 463
Van Buchem's disease, and sclerosteosis 414, 416
vas deferens, congenital bilateral absence (CBAVD) 155
vascular endothelial growth factor (VEGF) 498–9
velocardiofacial syndrome, and schizophrenia 462
visual handicap 488–504
  age-related cataract 488–93
  age-related macular degeneration 489, 496–501
  glaucoma 493–6
vitamin D
  deficiency 169
  and tuberculosis 287
vitamins, antioxidants 392–3
VLDL lipoproteins 362
  beta VLDL 365

Wright–Fisher model 47

Xenopus laevis, *xMeCP2* target gene 29
xeroderma pigmentosum 435

yeast two-hybrid screen 37